

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2006, 23:54:37 ; Search time 11385 Seconds
(without alignments)
11666.122 Million cell updates/sec

Title: SE01-458A
Perfect score: 2077
Sequence: 1 atcccggtcgaggattctct.....acctccccaaaactgttggt 2077

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : GenEmbl.*

- 1: gb_env.*
- 2: gb_pat.*
- 3: gb_ph.*
- 4: gb_pl.*
- 5: gb_pr.*
- 6: gb_ro.*
- 7: gb_scs.*
- 8: gb_sy.*
- 9: gb_un.*
- 10: gb_vi.*
- 11: gb_ov.*
- 12: gb_htg.*
- 13: gb_in.*
- 14: gb_om.*
- 15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	2075.4	99.9	2077	2	CS175961		CS175961 Sequence
2	2075.4	99.9	2077	5	AF420371		AF420371 Homo sapi
3	2056.4	99.0	2139	2	AX818172		AX818172 Sequence
4	2056.4	99.0	2139	5	AF061034		AF061034 Homo sapi
5	1989.4	95.8	2514	5	BC032762		BC032762 Homo sapi
6	1976.8	95.2	3454	2	CS071068		CS071068 Sequence
7	1927.4	92.8	2008	5	AF420373		AF420373 Homo sapi
8	1845.6	88.9	2464	5	AK055403		AK055403 Homo sapi
9	1779.4	85.7	1856	5	AF420372		AF420372 Homo sapi
10	1779.4	85.7	2318	5	BC013876		BC013876 Homo sapi
11	1747.8	84.2	2076	5	AF228374		AF228374 Macaca mu
12	1730.8	83.3	1734	2	CQ771153		CQ771153 Sequence
13	1672.6	80.5	2681	5	AB056409		AB056409 Macaca fa
14	1658.2	79.8	2671	5	AB063036		AB063036 Macaca fa
15	1629.4	78.4	1925	5	AV228373		AV228373 Macaca mu
16	1303	62.7	1930	14	BC102937		BC102937 Bos tauru
17	1291	62.2	1799	14	AF513722		AF513722 Sus scrof
18	1102.4	53.1	2443	6	BC061185		BC061185 Mus muscu

19	1101.8	53.0	2023	6	AY071834		AY071834 Mus muscu
20	1073.4	51.7	1716	6	AB222074		AB222074 Rattus no
21	1058.8	51.0	3577	6	AB069907		AB069907 Rattus no
22	997.6	48.0	1897	6	BC086976		BC086976 Rattus no
23	743.8	35.8	1163	2	CQ729566		CQ729566 Sequence
24	665.2	32.0	710	2	BD216639		BD216639 Novel hum
25	665.2	32.0	710	2	AR770229		AR770229 Sequence
26	645.2	31.1	3853	6	AB050777		AB050777 Rattus no
27	644.4	31.0	2476	11	AF389351		AF389351 Gallus ga
28	637.6	30.7	2555	11	AF380358		AF380358 Gallus ga
29	629.8	30.3	1496	6	AB222073		AB222073 Rattus no
30	620.2	29.9	1913	5	AK122746		AK122746 Homo sapi
31	526.6	25.4	3454	2	CS071068		CS071068 Sequence
32	458	22.1	458	2	AX777990		AX777990 Sequence
33	433.4	20.9	435	5	AF049614		AF049614 Homo sapi
34	424.6	20.4	444	2	AX969999		AX969999 Sequence
35	424.6	20.4	444	2	BD108718		BD108718 EST and e
36	424.6	20.4	444	2	AR413165		AR413165 Sequence
37	312	15.0	493	2	CQ923338		CQ923338 Sequence
38	312	15.0	493	2	CS135191		CS135191 Sequence
39	301.6	14.5	2301	11	BC088724		BC088724 Xenopus l
40	292.4	14.1	505	2	CS213387		CS213387 Sequence
41	292.4	14.1	505	2	CS217029		CS217029 Sequence
42	206	9.9	1203	2	CQ771155		CQ771155 Sequence
43	206	9.9	2555	5	AF283519S3		AF283519 Homo sapi
44	206	9.9	196203	5	AL355355		AL355355 Human DNA
45	197.2	9.5	206	2	CS070860		CS070860 Sequence
46	195.4	9.4	2536	11	BC063986		BC063986 Danio rer
47	193.2	9.3	257	5	AY749110S02		AY749110 Macaca mu
48	193	9.3	193	2	CQ923519		CQ923519 Sequence
49	193	9.3	193	2	CS135190		CS135190 Sequence
50	191.6	9.2	2075	11	BC055628		BC055628 Danio rer
51	191.4	9.2	240	2	CS175960		CS175960 Sequence
52	191.4	9.2	2566	5	AF283519S2		AF283520 Homo sapi
53	180.4	9.2	1166	2	CQ771154		CQ771154 Sequence
54	190.4	9.2	208202	12	AC013446		AC013446 Homo sapi
55	186	9.0	1183	2	CQ771156		CQ771156 Sequence
56	186	9.0	2875	5	AF283519S4		AF283522 Homo sapi
57	185.8	8.9	1587	11	CR523275		CR523275 Gallus ga
58	182.8	8.8	208202	12	AC013446		AC013446 Homo sapi
59	180.8	8.7	397	5	AY749110S01		AY749110 Macaca mu
60	177.2	8.5	307	5	AY749110S03		AY749112 Macaca mu
61	172.8	8.3	239066	12	AC105577		AC105577 Rattus no
62	162.4	7.8	1159	2	CQ771163		CQ771163 Sequence
63	162.4	7.8	5047	5	AF283519S7		AF283525 Homo sapi
64	158	7.6	469	5	AY749110S10		AY749119 Macaca mu
65	157	7.6	1122	2	CQ771166		CQ771166 Sequence
66	157	7.6	3575	5	AF283519S9		AF283527 Homo sapi
67	156.8	7.5	293	6	AY340623S02		AY340624 Mus muscu
68	156.8	7.5	116419	6	AL928662		AL928662 Mouse DNA
69	156.8	7.5	231611	12	AC104870		AC104870 Mus muscu
70	156.4	7.5	213772	12	AC153714		AC153714 Bos tauru
71	155.8	7.3	1153	2	CQ771158		CQ771158 Sequence
72	152.4	7.3	1150	2	CQ771161		CQ771161 Sequence
73	151	7.3	4015	5	AF283519S5		AF283523 Homo sapi
74	150.6	7.3	354	5	AY749110S05		AY749114 Macaca mu
75	146.4	7.0	340	5	AY749110S13		AY749122 Macaca mu
76	146	7.0	438	5	AY749110S08		AY749117 Macaca mu
77	135.4	6.5	660	6	AY340623S01		AY340623 Mus muscu
78	135.4	6.5	1131	2	CQ771164		CQ771164 Sequence
79	135.4	6.5	5021	5	AF283519S8		AF283526 Homo sapi
80	134.8	6.5	428	2	AX980154		AX980154 Sequence
81	134.8	6.5	428	2	BD115013		BD115013 EST and e
82	134.8	6.5	428	2	AR419460		AR419460 Sequence
83	130.6	6.3	512	5	AY749110S11		AY749120 Macaca mu
84	124	6.0	1094	2	CQ771162		CQ771162 Sequence
85	118	5.7	2139	5	AF283519S6		AF283524 Homo sapi
86	117.8	5.7	437	6	AY340623S10		AY340632 Mus muscu
87	116.4	5.5	1116	2	CQ771160		CQ771160 Sequence
88	114	5.5	267	6	AY340623S08		AY340630 Mus muscu
89	111.6	5.4	231	5	AY749110S07		AY749116 Macaca mu
90	107	5.2	1103	2	CQ771159		CQ771159 Sequence
91	106	5.1	379	6	AY340623S03		AY340625 Mus muscu

c	92	102.8	4.9	614	7	BV440176
	93	101.6	4.9	267	6	AY340623S01
	94	100.2	4.8	412	6	AY340623S11
	95	99	4.8	387	5	AY749111S0S06
	96	94.8	4.6	175	2	CQ920003
	97	93.6	4.5	2190	11	BC091725
	98	93.6	4.3	480	6	AY340623S05
	99	88.4	4.2	1080	2	CQ771165
	100	84	4.0	1500	2	CS031020
	101	84	4.0	1500	2	CS039972
102	84	4.0	1500	5	AF070533	
103	82.8	4.0	282	5	AY749111S0S12	
104	82.8	4.0	344	6	AY340623S13	
105	80.8	3.9	609	5	AY749111S0S09	
106	80.8	3.9	1828	14	BT020872	
107	80.8	3.9	2089	14	BT020767	
108	80.8	3.9	2130	14	BT020753	
109	80.8	3.9	2175	14	BT4414557	
110	80.8	3.9	2718	2	166494	
c	111	79	3.8	1074	2	CQ771157
	112	78.6	3.8	444	2	CQ398867
	113	78.6	3.8	444	2	CQ405147
	114	78	3.8	7303	5	AF283519S1
	115	77.8	3.7	1239	2	CS059891
	116	77.8	3.7	1748	6	AF069542
	117	77.8	3.7	1798	6	AF513109
	118	77.8	3.7	1874	2	AR454143
	119	77.8	3.7	2250	6	AY112937
	120	77.8	3.7	2282	6	BC021431
121	76.8	3.7	2138	14	BT020840	
122	76.6	3.7	1260	5	BT019621	
123	76.6	3.7	1260	8	AY888678	
124	76.6	3.7	1966	5	BC000299	
125	76.6	3.7	1975	5	AF062089	
126	76.6	3.7	1990	5	BC050612	
127	76.6	3.7	1994	5	AF074382	
128	76.6	3.7	2009	2	BD135431	
129	76.6	3.7	2009	2	CQ867779	
130	76.6	3.7	2009	2	AF534835	
131	76.6	3.7	2009	2	AX017214	
132	76.6	3.7	2034	2	BD135432	
133	76.6	3.7	2034	2	CQ867780	
134	76.6	3.7	2034	2	AR534836	
135	76.6	3.7	2034	2	AX017215	
136	76.6	3.7	2035	2	CS059897	
137	76.6	3.7	2035	2	AR609092	
138	76.6	3.7	2035	5	AF091453	
139	76.6	3.7	2071	5	BC012114	
140	76.6	3.7	2073	5	AX114157	
141	76.6	3.7	2121	5	AF261086	
142	76.2	3.7	1239	6	AY392762	
143	74.2	3.6	257	5	AY749111S0S04	
144	73	3.5	168119	6	AC118935	
145	69.2	3.3	1812	11	AY735398	
146	68.4	3.3	379	6	AY340623S09	
147	67.6	3.3	2480	11	BC091829	
148	66.8	3.2	2038	2	AX781112	
149	66.4	3.2	251418	12	AC169604	
c	150	65.2	3.1	27547	12	AC128056
	150	65.2	3.1	27547	12	AC128056

ALIGNMENTS

RESULT 1	LOCUS	SEQUENCE	LENGTH	DATE	LINEAR	PAT 12-OCT-2005
CS175961	DEFINITION	Sequence 113 from Patent WO2005090602.	2077 bp	DNA		
CS175961	ACCESSION	CS175961				
CS175961	VERSION	CS175961.1	GI:77620627			
CS175961	KEYWORDS	.				
CS175961	SOURCE	Homo sapiens (human)				
CS175961	ORGANISM	Homo sapiens				

[illegible]

Dd	2041	TTAAGTGTGATGATACCTCCCAAAACTGGTTGGT	2077
RESULT 3			
AX818172			
LOCUS	AX818172	2139 bp	DNA linear PAT 10-DEC-2003
DEFINITION	Sequence 43 from Patent WO03068268.		
ACCESSION	AX818172		
VERSION	AX818172.1	GI:39723242	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
JOURNAL	Hominidae; Homo.		
REFERENCE	1 Ek,S., Borrebaeck,C.A. and Ehinger,M.		
AUTHORS	Treatment, diagnosis and imaging of disease		
TITLE	Patent: WO 03068268-A 43 21-AUG-2003;		
JOURNAL	BIOINVENT INTERNATIONAL AB (SE)		
FEATURES	Location/Qualifiers		
source	1..2139		
ORIGIN	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query Match	99.0%; Score 2056.4; DB 2; Length 2139;		
Best Local Similarity	99.7%; Pred. No. 0;		
Matches 2071; Conservative	0; Mismatches 6; Indels 1; Gaps 1;		
Qy	1	ATCCCCGTCGGGAGTTCTCTCCAGCGGCACGATGCCGAGGAACAAGTGCACCTTGAGCGA	60
Dd	16	ATCCCCGTCGGGAGTTCTCTCCAGCGGCACGATGCCGAGGAACAAGTGCACCTTGAGCGA	75
Qy	61	AGCCAAGCCGGCGCAGGTGTGCTTTGATAGCTGTGTGTCACCTTCTGCGCTTGGGA	120
Dd	76	AGCCAAGCCGGCGCAGGTGTGCTTTGATAGCTGTGTGTCACCTTCTGCGCTTGGGA	135
Qy	121	TGAGCCGTACGCCCTCTGTAAACCACCTTCTCTCACCTTTGAAAACAGCTGCTGTGTCAGC	180
Dd	136	TGAGCCGTACGCCCTCTGTAAACCACCTTCTCTCACCTTTGAAAACAGCTGCTGTGTCAGC	195
Qy	181	ATTAAATGAAGAATTAGTCAGTGACAGGCGCTGGTGTGCTGAGTCCGCAATAGAGAATCAA	240
Dd	196	ATTAAATGAAGAATTAGTCAGTGACAGGCGCTGGTGTGCTGAGTCCGCAATAGAGAATCAA	255
Qy	241	AAATGTCAAAAATGTAAGTGGAGAGAAGTGGGCAACTTTTGGGAGTGAATTTTCCACAG	299
Dd	256	AAATGTCAAAAATGTAAGTGGAGAGAAGTGGGCAACTTTTGGGAGTGAATTTTCCACAG	315
Qy	300	GAACTTCTGCAATGTCCTCATCACTCTCAGCTGCCTCCTGAAAGAGGAGGAGAGCCCA	359
Dd	316	GAACTTCTGCAATGTCCTCATCACTCTCAGCTGCCTCCTGAAAGAGGAGGAGAGCCCA	375
Qy	360	GTGAAAGCAGAGAAATGAGACCCCACCTTGGCCCAACCCAAACCTTGAACTTTACCC	419
Dd	376	GTGAAAGCAGAGAAATGAGACCCCACCTTGGCCCAACCCAAACCTTGAACTTTACCC	435
Qy	420	CGGAGGAGCTGCTGCAGCAGATGAAAGAGCTCTTGACCAAGAACCCACAGCTGAAAGAAG	479
Dd	436	CGGAGGAGCTGCTGCAGCAGATGAAAGAGCTCTTGACCAAGAACCCACAGCTGAAAGAAG	495
Qy	480	CCATGAAGCTTAAATAATCAAGCCATGAAAGGGAGATTTGAGGAGCTTTCGGCCTGGACAG	539
Dd	496	CCATGAAGCTTAAATAATCAAGCCATGAAAGGGAGATTTGAGGAGCTTTCGGCCTGGACAG	555
Qy	540	AGAAACAGAGAGGAGAAAGCGCCAGTTTTTTTGAGATACAGAGCAAGAGCAAGAGAGCGTC	599
Dd	556	AGAAACAGAGAGGAGAAAGCGCCAGTTTTTTTGAGATACAGAGCAAGAGCAAGAGAGCGTC	615
Qy	600	TAATGGCCTTAGTTCATGAGAAATGAGAAATTTGAAGGAAGAGCTTGGAAACTTAAAAAGGGA	659

QY 1740 CTGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCATTGCAGCTGG 1799
Db 1756 CTGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCATTGCAGCTGG 1815
QY 1800 CAGTTCTGCTGAAGAGAGATGCTTTTCAAGCAGGAGCAGGAGCTTTCATGGAGA 1859
Db 1816 CAGTTCTGCTGAAGAGAGATGCTTTTCAAGCAGGAGCAGGAGCTTTCATGGAGA 1875
QY 1860 TGCAGAGTCTGATGGCGGAGAACAGTCACTCTGACAGCAGGCTTACCTTGTTCAAA 1919
Db 1876 TGCAGAGTCTGATGGCGGAGAACAGTCACTCTGACAGCAGGCTTACCTTGTTCAAA 1935
QY 1920 GAGGAGCTGAGGAGCAGGAGCTGGCGGCAACAGCGGAAATATTCGATTTCCTGCCCCA 1979
Db 1936 GAGGAGCTGAGGAGCAGGAGCTGGCGGCAACAGCGGAAATATTCGATTTCCTGCCCCA 1995
QY 1980 AGTGTGAGAGGTTTTCCTGATACATAGACAGCTTACGATTACGTTGATGATGATCA 2039
Db 1996 AGTGTGAGAGGTTTTCCTGATACATAGACAGCTTACGATTACGTTGATGATGATCA 2055
QY 2040 TTTAAGTGTGATGATACCTCCCAAACTGTGTGT 2077
Db 2056 TTTAAGTGTGATGATACCTCCCAAACTGTGTGT 2093

RESULT 4
AF061034 2139 bp mRNA linear PRI 09-MAY-1998
LOCUS Homo sapiens FIP2 alternatively translated mRNA, complete cds.
DEFINITION AF061034
ACCESSION AF061034
VERSION AF061034.1 GI:3127082
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2139)
Li, Y., Kang, J. and Horwitz, M.S.
AUTHORS Interaction of an adenovirus E3 14.7-kilodalton protein with a
TITLE novel tumor necrosis factor alpha-inducible cellular protein
containing leucine zipper domains
JOURNAL Mol. Cell. Biol. 18 (3), 1601-1610 (1998)
PUBMED 9488477
REFERENCE 2 (bases 1 to 2139)
Li, Y., Kang, J. and Horwitz, M.S.
AUTHORS Direct Submission
TITLE Submitted (21-APR-1998) Microbiology and Immunology, Albert
JOURNAL Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY
10461, USA

FEATURES
source 1..2139
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HeLa"
327..2060
/note="interacts with adenovirus E3-14.7kDa, a TNF-alpha
cytolysis antagonist; leucine zipper protein;
alternatively translated; long form"
/codon_start=1
/product="FIP2"
/protein_id="AAC16046.1"
/db_xref="GI:3127083"
/translation="MSHQPLSCLEKEDSPESNGNGPPHLAHPNLDFTPEELQOM
KELLETHQKANKLNQAMKGRPEELSAWTEKQKEROFFFIQSKERKERMALSH
ENEKLEKELGKLGKSSSEDPDSDRLPRAEAEQKDLQRTQVVRVLRQAEKADLGI
VSEQLKLNSSGSEDSFVIRMAEAGEAGSVKEIKHSPGSTRVTGTSALSHYRRS
ADGAKNTFEHEELTVSOLLCLREGNQKVERLVALEKAEKRVSDFEKKTNSRSEIEK
QTEGSTKEKDEKGPETVGSVEALNLQVTSLEKELAEKALASKOLOMDENKQ
QALERKNSAIPSELNEKQELVYPNKKLEQVSEMLSEIKMEQAKTEDEKSKLTVLQWT
HNKLLQEHNNALKTIELTRKSEKVDRAVLKELSEKLEAEKALASKOLOMDENKQ

CDS

CDS

IAKQEDLETWTLIRAQNEVYCSDFHAEPAAREKIHKEEQLALQALQVLLKENDAPED
GGRSLMEMQSRHGARTSDSDQAYLVORGAEDRWROQRNIPIHSCPKGEVLDPID
TLQIHVMDCII"
498..2060
/note="interacts with adenovirus E3-14.7kDa, a TNF-alpha
cytolysis antagonist; leucine zipper protein;
alternatively translated; short form"
/codon_start=1
/product="FIP2"
/protein_id="AAC16047.1"
/db_xref="GI:3127084"
/translation="MKLNQAMKGRPEELSAWTEKQKEROFFFIQSKERKERMALSH
HNEKLEKELGKLGKSSSEDPDSDRLPRAEAEQKDLQRTQVVRVLRQAEKADLGI
VSEQLKLNSSGSEDSFVIRMAEAGEAGSVKEIKHSPGSTRVTGTSALSHYRRR
SADGAKNTFEHEELTVSOLLCLREGNQKVERLVALEKAEKRVSDFEKKTNSRSEIEK
QTEGSTKEKDEKGPETVGSVEALNLQVTSLEKELAEKALASKOLOMDENKQ
COALERKNSAIPSELNEKQELVYPNKKLEQVSEMLSEIKMEQAKTEDEKSKLTVLQOM
HNKLLQEHNNALKTIELTRKSEKVDRAVLKELSEKLEAEKALASKOLOMDENKQ
TTAKQEDLETWTLIRAQNEVYCSDFHAEPAAREKIHKEEQLALQALQVLLKENDAPED
DGRSLMEMQSRHGARTSDSDQAYLVORGAEDRWROQRNIPIHSCPKGEVLDPID
DTLQIHVMDCII"

ORIGIN

Query Match 99.0%; Score 2056.4; DB 5; Length 2139;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2071; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1 ATCCCGGTGGGAGTTCTCTCCAGGCGGACGATGCGGAGGAAACAGTACCTGAGCGA 60
Db 16 ATCCCGGTGGGAGTTCTCTCCAGGCGGACGATGCGGAGGAAACAGTACCTGAGCGA 75
QY 61 AGCCAAAGCCGGCGGAGGTGGCTTTGATAGCTGGTGGTGCACCTTCCTGGCCTTGA 120
Db 76 AGCCAAAGCCGGCGGAGGTGGCTTTGATAGCTGGTGGTGCACCTTCCTGGCCTTGA 135
QY 121 TGAGCCGTACGCTCTGTAAACCCCACTTCCTCACCTTTGAAACAGCTGCTGTTCCAGC 180
Db 136 TGAGCCGTACGCTCTGTAAACCCCACTTCCTCACCTTTGAAACAGCTGCTGTTCCAGC 195
QY 181 ATTAATGAAGATTAGTCAGTCAGCAGCGCTGGTGTGCTGAGTCGCCACATAGAAGATCAA 240
Db 196 ATTAATGAAGATTAGTCAGTCAGCAGCGCTGGTGTGCTGAGTCGCCACATAGAAGATCAA 255
QY 241 AAATGTCAAAATGTAACCTGGAGAGAAAGTGGGCAACTTTT-GGAGTGACTTTTCCACAG 299
Db 256 AAATGTCAAAATGTAACCTGGAGAGAAAGTGGGCAACTTTTGGGAGTGACTTTTCCACAG 315
QY 300 GAATCTCTCAATGTCCCATCAACCTCTCAGCTGCTCCTCACTGAAAAGAGGAGCAGCCCCA 359
Db 316 GAATCTCTCAATGTCCCATCAACCTCTCAGCTGCTCCTCACTGAAAAGAGGAGCAGCCCCA 375
QY 360 GTGAAAGCAGAGAAATGGACCCCCCACTGGCCCAACCCCAACCTGGGACACGTTTACCC 419
Db 376 GTGAAAGCAGAGAAATGGACCCCCCACTGGCCCAACCCCAACCTGGGACACGTTTACCC 435
QY 420 CGAGAGAGCTGTCAGCAGATGAAGAGCTCTGACCAAGAACACCACCTGAAAAGAG 479
Db 436 CGAGAGAGCTGTCGACGAGATGAAGAGCTCTGACCCGAGAACCCACCTGAAAAGAG 495
QY 480 CCATGAGCTTAATAATCAAGCCATGAAGGAGATTTTGAAGGAGCTTCGCGCTGGACAG 539
Db 496 CCATGAGCTTAATAATCAAGCCATGAAGGAGATTTTGAAGGAGCTTCGCGCTGGACAG 555
QY 540 AGAAACAGAGGAGAAACGCCAGTTTTTTTGAGATACAGACAAAGAGCAAGAGAGCGTC 599
Db 556 AGAAACAGAGGAGAAACGCCAGTTTTTTTGAGATACAGACAAAGAGCAAGAGAGCGTC 615
QY 600 TAATGGCTTGAATCATGAGAAATGAGAAATTTGAGGAGAGCTTGGAAACTAAAGGGA 659
Db 616 TAATGGCTTGAATCATGAGAAATTTGAGGAGAGAGCTTGGAAACTAAAGGGA 675
QY 660 AATCAGAAAGGTCATCTGAGGAGCCCACTGATGATCTCCAGGCTTCCAGGCGCCGAAGCGG 719

Db 676 AATCAGAAAGGTATCTGAGGACCCCACTGATGACTCCAGGCTTCCACGGCCGAGCGG 735
Qy 720 AGCAGGAAAGGACCCAGCTCAGGACCCAGGTGGTGGAGCTTACAAGCAGAGAAGGCGAGACC 779
Db 736 AGCAGGAAAGGACCCAGCTCAGGACCCAGGTGGTGGAGCTTACAAGCAGAGAAGGCGAGACC 795
Qy 780 TGTGGGCAATCGTGTCTGAATCTGAGCTCAAGCTGAACTCAGCGGCTTCTCAGAGATT 839
Db 796 TGTGGGCAATCGTGTCTGAATCTGAGCTCAAGCTGAACTCAGCGGCTTCTCAGAGATT 855
Qy 840 CCTTTGTTGAAATTAGGATGCTGAAGGAGAGACGAAGGCTCAGTAAAGCAAAATCAAGC 899
Db 856 CCTTTGTTGAAATTAGGATGCTGAAGGAGAGACGAAGGCTCAGTAAAGCAAAATCAAGC 915
Qy 900 ATAGTCTTGGGCCCAACGAGACAAGTCTCACTGGCAGCGCAATGTCTTAAATATAGGAGCA 959
Db 916 ATAGTCTTGGGCTCCACGAGAACAGTCTCACTGGCAGCGCAATGTCTTAAATATAGGAGCA 975
Qy 960 GATCTGCAGATGGGGCCCAAGAAATTAATCTCGAACATGAGGAGTTAACTGTGAGCCAGCTCC 1019
Db 976 GATCTGCAGATGGGGCCCAAGAAATTAATCTCGAACATGAGGAGTTAACTGTGAGCCAGCTCC 1035
Qy 1020 TGCTGTGCTTAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGGAGG 1079
Db 1036 TGCTGTGCTTAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGGAGG 1095
Qy 1080 CCAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAAATCGTTCTGAGATTGAAACCC 1139
Db 1096 CCAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAAATCGTTCTGAGATTGAAACCC 1155
Qy 1140 AGACAGAGGGAGACACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTTGTGAA 1199
Db 1156 AGACAGAGGGAGACACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTTGTGAA 1215
Qy 1200 GCGAAGTGGAGCACTGAACCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTC 1259
Db 1216 GCGAAGTGGAGCACTGAACCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTC 1275
Qy 1260 ATACAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAGGCC 1319
Db 1276 ATACAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAGGCC 1335
Qy 1320 TTGAAGGAAATTTCTGCAATTTCCATCAGAGTTGAATGAAGAAAGCAAGCTTGTGTTATA 1379
Db 1336 TTGAAGGAAATTTCTGCAATTTCCATCAGAGTTGAATGAAGAAAGCAAGCTTGTGTTATC 1395
Qy 1380 CTACAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGG 1439
Db 1396 CTACAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGG 1455
Qy 1440 CTAACACAGAGGATGAAAGTCCAAATTTAACTGTGCTACAGATGACACACAAAGCTTC 1499
Db 1456 CTAACACAGAGGATGAAAGTCCAAATTTAACTGTGCTACAGATGACACACAAAGCTTC 1515
Qy 1500 TTCAAGAACATAATGCAATTGAAACAAATTTGAGGAACATAACAGAAAGAGTTCAGAAA 1559
Db 1516 TTCAAGAACATAATGCAATTGAAACAAATTTGAGGAACATAACAGAAAGAGTTCAGAAA 1575
Qy 1560 AAGTGGACAGGCGAGTCTGAAGCACTGAGTGGAAAGCACTGGAAGTGGCAGAGAGCTC 1619
Db 1576 AAGTGGACAGGCGAGTCTGAAGCACTGAGTGGAAAGCACTGGAAGTGGCAGAGAGCTC 1635
Qy 1620 TGGCTTCCAAACAGCTGCAATGGAATGAAGCAAAACCATTTGCCAAGCAGGAAGG 1679
Db 1636 TGGCTTCCAAACAGCTGCAATGGAATGAAGCAAAACCATTTGCCAAGCAGGAAGG 1695
Qy 1680 ACTTGGAAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTTCAATTTTCATG 1739
Db 1696 ACTTGGAAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTTCAATTTTCATG 1755
Qy 1740 CTGAAGACAGCGAGAGAGAAATTCATGAGGAAAGAGCAACTGGCATTTGAGCTGG 1799
Db 1756 CTGAAGACAGCGAGAGAGAAATTCATGAGGAAAGAGCAACTGGCATTTGAGCTGG 1815

Qy 1800 CAGTTCTGCTGAAAAGAGAAATGATGCTTTTTCGAAGACGAGGCGAGGCTCTTGTATGGAGA 1859
Db 1816 CAGTTCTGCTGAAAAGAGAAATGATGCTTTTTCGAAGACGAGGCGAGGCTCTTGTATGGAGA 1875
Qy 1860 TGCAGAGTCTGTCATGGGGCGAGAAACAAGTGAATCTGACCAAGAGGCTTACCTTGTTCAAA 1919
Db 1876 TGCAGAGTCTGTCATGGGGCGAGAAACAAGTGAATCTGACCAAGAGGCTTACCTTGTTCAAA 1935
Qy 1920 GAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTCAGATCAATTCCTGCCCCA 1979
Db 1936 GAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTCAGATCAATTCCTGCCCCA 1995
Qy 1980 AGTGTGAGAGGTTTCTGCTTGACATAGACAGCTTACAGATTCAGCTGATGATTCATCA 2039
Db 1996 AGTGTGAGAGGTTTCTGCTTGACATAGACAGCTTACAGATTCAGCTGATGATTCATCA 2055
Qy 2040 TTTAAGTGTGATGATATACCTCCCAAAACTGTTGGT 2077
Db 2056 TTTAAGTGTGATGATATACCTCCCAAAACTGTTGGT 2093

RESULT 5

BC032762

LOCUS

DEFINITION

Homo sapiens optineurin, mRNA (cDNA clone MGC:44991 IMAGE:3457195),

complete cds.

ACCESSION

BC032762

VERSION

BC032762.1

KEYWORDS

MGC.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1 (bases 1 to 2514)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, K.H., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Amerson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Jones, S.J., Skalska, U., Smal, D.E.,

Schnerch, A., Schein, J.E., Jones, S.J., Skalska, U., Smal, D.E.,

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

PUBMED

2 (bases 1 to 2514)

REFERENCE

Strausberg, R.

Direct Submission

Submitted (06-JUN-2002) National Institutes of Health, Mammalian

Gene Collection (MGC) Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
 Series: IRAK Plate: 68 Row: a Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 20149571.

FEATURES

source
 1. .2514
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:44991 IMAGE:3457195"
 /tissue_type="Cervix, carcinoma"
 /clone_lib="NIH_MGC_12"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1. .2514
 /gene="OPTN"
 /note="synonyms: FIP2, HYPL, TFIIIA-INTP, NRP, GLC1E,
 HIP7"
 /db_xref="GeneID:10133"
 /db_xref="MIM:602432"
 359..2074
 /gene="OPTN"
 /codon_start=1
 /product="OPTN protein"
 /protein_id="AAH32762.1"
 /db_xref="GI:21619683"
 /db_xref="GeneID:10133"
 /db_xref="MIM:602432"
 /translation="MSHQPLSCITEKEDSPSESTNGPPHLHPNLDTFPEELLQOM
 KELLTENHQLKEAKLNQAMKGRFEELSAWTKQKEEROFFEIQSKEAKERLWALSH
 ENKLEKELGKJGKSRSSEDDDSRLPRAEAEQKDLQRTQVRLQAEKADLIGI
 VSELQKLNSGSDSFVEIRMAEGEAGSVKEIKHSPGTRTVSTGSRSDAGKN
 YFEHEETVSOLLCLREGNOKVERLEVALKEKRVSDPEKKTNSRSEITOTEGST
 EKENDEKGPETVGSVEALNLOVSLFKELQEAHTKLSAEIMKKELQSKCOALERK
 NSAIPSELNKEQLVYTNKKLEQVESNLSEIKMEQAKTEDESKLTVLQMTNKLQ
 EHNALXTIELTRKSEKVDRAVLKELSEKLEAEKALASKOLOMDKQTIKQEE
 DLEMTITLRAQMEVYCSDFHAERAAEKIHEEKELQALQALVLLKENDAFEDGGROSL
 MEMOSRHGARTSDSDQALVLRGAEDRDWRQORNPPIHSCPKGGEVLPIDITLIQIHV
 MDCII"
 1055..1588
 /gene="OPTN"
 /note="Smc; Region: Chromosome segregation ATPases [Cell
 division and chromosome partitioning]"
 /db_xref="CDD:COG1196"

misc_feature
 95.8%; Score 1989.4; DB 5; Length 2514;
 Query Match
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 2043; Conservative 0; Mismatches 1; Indels 33; Gaps 2;

ORIGIN

Db 184 TGAGCGGTACGCCTCTGTAAACCCCAACTTCCTCACCTTTGAAACAGCTGCTGGTTTCAGC 243

Qy 181 ATTAATGAAGATTAGTCAGTGACAGGCTGTGGTGTGCTAGTCCGCACATAGAAGAAATCAA 240

Db 244 ATTAATGAAGATTAGTCAGTGACAGGCTGTGGTGTGCTAGTCCGCACATAGAAGAAATCAA 303

Qy 241 AAATGTCCAAAATGTAATCGGAGAGAAAGTGGCAACTTTTGGAGTGACTTTTCCACAGG 300

Db 304 AAATGTCCAAAATGTAATCGGAGAGAAAGTGGCAACTTTTGGAGTGACTTTTCCACAGG 348

Qy 301 AACTTCTGCAATGTCCCACTCAACCTCTCAGCTGCCTCACTGAAAAGGAGACAGCCCCAG 360

Db 349 AACTTCTGCAATGTCCCACTCAACCTCTCAGCTGCCTCACTGAAAAGGAGACAGCCCCAG 408

Qy 361 TGAAGCACAGGAAATGGACCCCACTGGGCCCACTGAGGAGCTTTTGGAGTGTACCTTACCCC 420

Db 409 TGAAGCACAGGAAATGGACCCCACTGGGCCCACTGAGGAGCTTTTGGAGTGTACCTTACCCC 468

Qy 421 GGAGGAGCTGCTGCAGCAGATGAAGAGCTCTCTGACCAAGNACCACTGCTGAAAAGGAGC 480

Db 469 GGAGGAGCTGCTGCAGCAGATGAAGAGCTCTCTGACCGAGAACCACTGAAAAGGAGC 528

Qy 481 CATGAAGCTTAAATAATCAAGCCATGAAGGAGATTGTGAGGAGCTTTTGGAGTGTACCTTACCCC 540

Db 529 CATGAAGCTTAAATAATCAAGCCATGAAGGAGATTGTGAGGAGCTTTTGGAGTGTACCTTACCCC 588

Qy 541 GAAACAGAGAGGAGAACGCCAGTTTTTGGAGATACAGAGCAAGAACCAAGAGCGGTCT 600

Db 589 GAAACAGAGAGGAGAACGCCAGTTTTTGGAGATACAGAGCAAGAACCAAGAGCGGTCT 648

Qy 601 AATGGCCTTGTAGTCATGAGATGAGAAATGAAGGAGAGCTTTGGAATACTAAAGGGAA 660

Db 649 AATGGCCTTGTAGTCATGAGATGAGAAATGAAGGAGAGCTTTGGAATACTAAAGGGAA 708

Qy 661 ATCAGAAAGGTCACTGAGGAGCCCACTGATGACTCCAGGCTTCCAGGCGCCGAAAGCGGA 720

Db 709 ATCAGAAAGGTCACTGAGGAGCCCACTGATGACTCCAGGCTTCCAGGCGCCGAAAGCGGA 768

Qy 721 GCAGGAAAAGGACCACTCAGGACCCAGGCTGAGGCTCAAGAGAGAGAGGAGGAGCCT 780

Db 769 GCAGGAAAAGGACCACTCAGGACCCAGGCTGAGGCTCAAGAGAGAGAGGAGGAGCCT 828

Qy 781 GTTGGCATCGTCTCTGAATGAGCTCAAGCTCAAGCTCAAGCTCAAGCGGCTCCTCAGAAGATTC 840

Db 829 GTTGGCATCGTCTCTGAATGAGCTCAAGCTCAAGCTCAAGCGGCTCCTCAGAAGATTC 888

Qy 841 CTTTGTGAAAATAGCATGGCTGAAGAGAGAGAGGCTCAGTAAAGAAATCAAGCA 900

Db 889 CTTTGTGAAAATAGCATGGCTGAAGAGAGAGAGGCTCAGTAAAGAAATCAAGCA 948

Qy 901 TACTCTCTGGGCCCAACAGAACTCTCCACTGCGACGGCAATGCTTAAATATAGGAGCAG 960

Db 949 TACTCTCTGGGCCCAACAGAACTCTCCACTGCGAC-----GAGCAG 990

Qy 961 ATCTGACATGGGGCCCAAGAAATTAATCTCGAACATGAGGAGTAACTCTGAGCCAGCTCCT 1020

Db 991 ATCTGACATGGGGCCCAAGAAATTAATCTCGAACATGAGGAGTAACTCTGAGCCAGCTCCT 1050

Qy 1021 GCTGTGCTCAAGGAAAGGGAATCAGAGGTGGAGAGACTTTGAAGTTTGCATCTCAAGGAGGC 1080

Db 1051 GCTGTGCTCAAGGAAAGGGAATCAGAGGTGGAGAGACTTTGAAGTTTGCATCTCAAGGAGGC 1110

Qy 1081 CAAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAAAGTAATCGTCTGAGATTTGAAACCCA 1140

Db 1111 CAAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAAAGTAATCGTCTGAGATTTGAAACCCA 1170

Qy 1141 GACAGAGGAGACACAGAGAAAGAGATGATGAAGAGAGAGGCGCCGAGACTGTTTGAAG 1200

Db 1171 GACAGAGGAGACACAGAGAAAGAGATGATGAAGAGAGAGGCGCCGAGACTGTTTGAAG 1230

Qy 1201 CGAAGTGAAGACACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGGCTTCAAGAGGCTCA 1260

Db 1231 CGAAGTGAAGACACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGGCTTCAAGAGGCTCA 1290

Qy	1261	TACAAACTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAAGTGTCCAGGCCCT	1320
Db	1291	TACAAACTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAAGTGTCCAGGCCCT	1350
Qy	1321	TGAAGGAAAAATTTCTGCAATTCATCAGAGTTGATGAAAGCAAGAGCTTTTATAC	1380
Db	1351	TGAAGGAAAAATTTCTGCAATTCATCAGAGTTGATGAAAGCAAGAGCTTTTATAC	1410
Qy	1381	TACAAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGC	1440
Db	1411	TACAAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGC	1470
Qy	1441	TAAACAGAGGATGAAAGTCCAAATTAATCTGTCTACAGATGACACAAAGCTTCT	1500
Db	1471	TAAACAGAGGATGAAAGTCCAAATTAATCTGTCTACAGATGACACAAAGCTTCT	1530
Qy	1501	TCAAGACATAATATGATGAAACAAATTTGAGGAACTAACAGAAAGAGTCAAGAAA	1560
Db	1531	TCAAGACATAATATGATGAAACAAATTTGAGGAACTAACAGAAAGAGTCAAGAAA	1590
Qy	1561	AGTGGACAGGGCAGTGTGAAAGGAACCTGAGTGAAGAACTGGCAGAGAGGCTCT	1620
Db	1591	AGTGGACAGGGCAGTGTGAAAGGAACCTGAGTGAAGAACTGGCAGAGAGGCTCT	1650
Qy	1621	GGCTTCCAAACAGCTGCAAAATGGAATGAATGAAGCAAACTGCAAGCAGGAAGAGGA	1680
Db	1651	GGCTTCCAAACAGCTGCAAAATGGAATGAATGAAGCAAACTGCAAGCAGGAAGAGGA	1710
Qy	1681	CCTGGAACCATGACCTCCTCAGGGCTCAGATGGAAATTTACTGTTCTGATTTTCATGC	1740
Db	1711	CCTGGAACCATGACCTCCTCAGGGCTCAGATGGAAATTTACTGTTCTGATTTTCATGC	1770
Qy	1741	TGAAGAGCAGCAGAGAGAGAAAATTCATGAGAAAAGGAGCAACTGGCATTGCACTGGC	1800
Db	1771	TGAAGAGCAGCAGAGAGAGAAAATTCATGAGAAAAGGAGCAACTGGCATTGCACTGGC	1830
Qy	1801	AGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAGGAGGAGCTTGTGATGGAGAT	1860
Db	1831	AGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAGGAGGAGCTTGTGATGGAGAT	1890
Qy	1861	GCAGAGTCGTATGCGGCGAGAACAGTGAATCTGACAGGAGGAGGAGCTTGTGATGGAGAT	1920
Db	1891	GCAGAGTCGTATGCGGCGAGAACAGTGAATCTGACAGGAGGAGGAGCTTGTGATGGAGAT	1950
Qy	1921	AGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTCCGATTTCATTCGCCCCAA	1980
Db	1951	AGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTCCGATTTCATTCGCCCCAA	2010
Qy	1981	GTGTGGAGAGGTTCTGCTGACATAGACAGTTACAGATTCACGTGATGGATTGCAATCAT	2040
Db	2011	GTGTGGAGAGGTTCTGCTGACATAGACAGTTACAGATTCACGTGATGGATTGCAATCAT	2070
Qy	2041	TTAAGTGTGATGATATCACCTCCCAAACTGTTGGT	2077
Db	2071	TTAAGTGTGATGATATCACCTCCCAAACTGTTGGT	2107
RESULT 6			
LOCUS	CS071068	3454 bp	DNA linear PAT 05-MAY-2005
DEFINITION	Sequence 231 from Patent WO2001032927.		
ACCESSION	CS071068		
VERSION	CS071068.1 GI:63088472		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Sornasse,T., Seilhamer,J.J. and Watson,G.A.		
TITLE	Tissue specific genes of diagnostic import		

JOURNAL	Patent: WO 2001032927-A 231 10-MAY-2001;		
FEATURES	Incyte Genomics, Inc. (US)		
source	1. 3454		
ORIGIN	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query Match	95.28;	Score 1976.8;	DB 2; Length 3454;
Best Local Similarity	99.88;	Pred. No. 0;	
Matches 2000; Conservative	0;	Mismatches	2; Indels
	2;	Gaps	2;
Qy	1	ATCCCGGTGCGGAGTTCCTCTCCAGCGCGCACGATGCCGAGGAAACAGTGACCTTGAGCGA	60
Db	16	ATCCCGGTGCGGAGTTCCTCTCCAGCGCGCACGATGCCGAGGAAACAGTGACCTTGAGCGA	75
Qy	61	AGCCAAAGCCGGGCGGCGAGGTGCTTTGATAGCTGTGTGTGTCACCTTCCTGGCCTTGA	120
Db	76	AGCCAAAGCCGGGCGGCGAGGTGCTTTGATAGCTGTGTGTGTCACCTTCCTGGCCTTGA	135
Qy	121	TGAGCCCTACGCTCTGTAAACCCAACTTCTCACCTTTGAAACAGCTGCTGCTTCA	180
Db	136	TGAGCCCTACGCTCTGTAAACCCAACTTCTCACCTTTGAAACAGCTGCTGCTTCA	195
Qy	181	ATTAATGAAGATTAGTCAGTGACAGGCGCTGTGTGTGCTGAGTCCGCAATAGAAAGTCAA	240
Db	196	ATTAATGAAGATTAGTCAGTGACAGGCGCTGTGTGTGCTGAGTCCGCAATAGAAAGTCAA	255
Qy	241	AAATGTCCAAAATGTAACTGAGAGAAAGTGGGCAACTTTT-GGAGTGACTTTTCCACAG	299
Db	256	AAATGTCCAAAATGTAACTGAGAGAAAGTGGGCAACTTTTGGAGTGACTTTTCCACAG	315
Qy	300	GAACTTCTGCAATGTCCTCACTCAACCTCTCAGCTGCCTCACTGAAAGAGGACAGCCCA	359
Db	316	GAACTTCTGCAATGTCCTCACTCAACCTCTCAGCTGCCTCACTGAAAGAGGACAGCCCA	375
Qy	360	GTGAAAGCAGAGAAATGGAACCCCTGCGCCACCCCAACCT- GGCACAGCTTTTACC	418
Db	376	GTGAAAGCAGAGAAATGGAACCCCTGCGCCACCCCAACCTGCGGACAGCTTTTACC	435
Qy	419	CCGAGGAGCTGCTGAGCAGATGAAAGAGCTCTCTGACCAAGAACACCACTGAAAGAA	478
Db	436	CCGAGGAGCTGCTGAGCAGATGAAAGAGCTCTCTGACCGAGAACACCACTGAAAGAA	495
Qy	479	GCCATGAAGCTAATTAATCAAGCCATGAAAGGAGATTTGAGGAGCTTTCCGCTTGACA	538
Db	496	GCCATGAAGCTAATTAATCAAGCCATGAAAGGAGATTTGAGGAGCTTTCCGCTTGACA	555
Qy	539	GAGAAACAGAGGAAGAACGCCAGTTTGTGAGATACAGAGCAAGAAAGCAAGAGCGCT	598
Db	556	GAGAAACAGAGGAAGAACGCCAGTTTGTGAGATACAGAGCAAGAAAGCAAGAGCGCT	615
Qy	599	CTAATGCGCTTGAGTCTATGAGAAATGAGAAATTAAGGAGAGCTTTGAGAACTAAAGGG	658
Db	616	CTAATGCGCTTGAGTCTATGAGAAATGAGAAATTAAGGAGAGCTTTGAGAACTAAAGGG	675
Qy	659	AAATCAGAAAGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGCGCGAAGCG	718
Db	676	AAATCAGAAAGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGCGCGAAGCG	735
Qy	719	GAGCAGGAAAGGACAGCTCAGGACCCAGGCTGAGGCTTACAGCAGAGAGGAGGAGAC	778
Db	736	GAGCAGGAAAGGACAGCTCAGGACCCAGGCTCAGGACCCAGGCTGAGGCTTACAGCAGAGAGGAGAC	795
Qy	779	CTGTTGGCATCTGTCTGAACCTGACGCTCAAGCTGAACCTCCAGCGGCTCTCTCAGAGAT	838
Db	796	CTGTTGGCATCTGTCTGAACCTGACGCTCAAGCTGAACCTCCAGCGGCTCTCTCAGAGAT	855
Qy	839	TCCTTTGTTCAATTTAGGATGCTGAGGAGAGCAGAGGGTTCAGTTAAAGAAATCAAG	898
Db	856	TCCTTTGTTCAATTTAGGATGCTGAGGAGAGCAGAGGGTTCAGTTAAAGAAATCAAG	915


```
Qy 121 TGAGCCGTACGCCCTCTGTGTAACCCAACTTCTCACTTTGAAACAGCTGCTGGTTTCAGC 180
Db 121 TGAGCCGTACGCCCTCTGTGTAACCCAACTTCTCACTTTGAAACAGCTGCTGGTTTCAGC 180
Qy 181 ATTAATGAAGATTAGTCAGTGACAGGCTGGTGCTGAGTCCGCACATAGAGAATCAA 240
Db 181 ATTAATGAAGATTAGTCAGTGACAGGCTGGTGCTGAGTCCGCACATA----- 230
Qy 241 AAATGTCCAAAATGTAATCGAGAGAAAAGTGGGCAACTTTTGGAGTGACTTTTCCACAGG 300
Db 231 -----G 231
Qy 301 AACTTCTGCAATGCCATCAACTCTCAGCTGCTCACTGAAAGAGGACAGCCCCAG 360
Db 232 AACTTCTGCAATGCCATCAACTCTCAGCTGCTCACTGAAAGAGGACAGCCCCAG 291
Qy 361 TGAAGCACAGGAATGGACCCGCCACCTGGCCCAACCCAACTGACACAGTTTACCCC 420
Db 292 TGAAGCACAGGAATGGACCCGCCACCTGGCCCAACCCAACTGACACAGTTTACCCC 351
Qy 421 GGAGGAGCTGTGTCAGCAGATGAAGAGCTCTGACCAAGAACCAACAGCTGAAAGAACG 480
Db 352 GGAGGAGCTGTGTCAGCAGATGAAGAGCTCTGACCCAGAACCCACAGCTGAAAGAACG 411
Qy 481 CATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTTGAGGAGCTTTGCGCCTGGACAGA 540
Db 412 CATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTTGAGGAGCTTTGCGCCTGGACAGA 471
Qy 541 GAAACAGAGGAGAACGCCAGTTTTTTTGATACAGAGCAAGAACCAAAAGAGCGTCT 600
Db 472 GAAACAGAGGAGAACGCCAGTTTTTTTGATACAGAGCAAGAACCAAAAGAGCGTCT 531
Qy 601 AATGGCCCTCAGTCATCAGAAATGAGAAATTTGAAGGAGAGCTTGGAATCTAAAGGGAA 660
Db 532 AATGGCCCTCAGTCATCAGAAATGAGAAATTTGAAGGAGAGCTTGGAATCTAAAGGGAA 591
Qy 661 ATCAGAAAGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCCAAGCGGA 720
Db 592 ATCAGAAAGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCCAAGCGGA 651
Qy 721 GCAGGAAAGGACCCAGCTCAGGACCCAGGTGGTAGGCTACAGCAGAGAGGACGACCT 780
Db 652 GCAGGAAAGGACCCAGCTCAGGACCCAGGTGGTAGGCTACAGCAGAGAGGACGACCT 711
Qy 781 GTTGGGCATCGTGTCTGAACTGACGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAGATTC 840
Db 712 GTTGGGCATCGTGTCTGAACTGACGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAGATTC 771
Qy 841 CTTTGTGTAATAGGATGGCTGAAGGAGAGCAGAGGGTTCAGTAAAGAAATCAAGCA 900
Db 772 CTTTGTGTAATAGGATGGCTGAAGGAGAGCAGAGGGTTCAGTAAAGAAATCAAGCA 831
Qy 901 TAGTCTGGGCCCCACGAGACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGAGCAG 960
Db 832 TAGTCTGGGCCCCACGAGACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGAGCAG 891
Qy 961 ATCTCAGATGGGGCCCAAGAAATTAATTCGAACATGAGGAGTTAACTGTGAGCCAGCTCCT 1020
Db 892 ATCTCAGATGGGGCCCAAGAAATTAATTCGAACATGAGGAGTTAACTGTGAGCCAGCTCCT 951
Qy 1021 GCTGTGCTTAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTCACCTCAAGGAGGC 1080
Db 952 GCTGTGCTTAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTCACCTCAAGGAGGC 1011
Qy 1081 CAAAGAAAGATTTCAGATTTGAAAGAGAAAAAAGTAATCGTTCTGAGATTGAAACCCCA 1140
Db 1012 CAAAGAAAGATTTCAGATTTGAAAGAGAAAAAAGTAATCGTTCTGAGATTGAAACCCCA 1071
Qy 1141 GACAGGGGAGCACAGAGAGAGATGATGAGAGAAAGGCCCGGAGACTTTTGAAG 1200
Db 1072 GACAGGGGAGCACAGAGAGAGATGATGAGAGAAAGGCCCGGAGACTTTTGAAG 1131
```

```
Qy 1201 CGAAGTGGAAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCA 1260
Db 1132 CGAAGTGGAAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCA 1191
Qy 1261 TACAAAACCTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAGTGTTCAGGCCCT 1320
Db 1192 TACAAAACCTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAGTGTTCAGGCCCT 1251
Qy 1321 TGAAGGAAAAAATTCCTGCAATTCATCAGAGTTGAAATGAAAAGCAAGAGCTTGTTTATAC 1380
Db 1252 TGAAGGAAAAAATTCCTGCAATTCATCAGAGTTGAAATGAAAAGCAAGAGCTTGTTTATAC 1311
Qy 1381 TACAAAAGTGTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAAATGGAACAGGC 1440
Db 1312 TACAAAAGTGTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAAATGGAACAGGC 1371
Qy 1441 TAAAACAGAGGATGAAAAGTCCAAAATTAAGTGTGCTACAGATGACACACAACAGCTTCT 1500
Db 1372 TAAAACAGAGGATGAAAAGTCCAAAATTAAGTGTGCTACAGATGACACACAACAGCTTCT 1431
Qy 1501 TCAAGAACATTAATAATGTCATTGAAAACAAATTGAGGAACCTAACAGAAAAGAGTCAGAAA 1560
Db 1432 TCAAGAACATTAATAATGTCATTGAAAACAAATTGAGGAACCTAACAGAAAAGAGTCAGAAA 1491
Qy 1561 AGTGGACAGGCGAGTCTGTAAGGAACCTGAGTGA AAAAATCTGGAACCTGGCAGAGAGGCTCT 1620
Db 1492 AGTGGACAGGCGAGTCTGTAAGGAACCTGTAAGGAACCTGGAACCTGGCAGAGAGGCTCT 1551
Qy 1621 GGTCTTCAAAACAGCTGCAAAATGGAATGAAAGCAAAACCATTTGCCAAGCAGGAAGAGGA 1680
Db 1552 GGTCTTCAAAACAGCTGCAAAATGGAATGAAAGCAAAACCATTTGCCAAGCAGGAAGAGGA 1611
Qy 1681 CCTGGAAAACCATCACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTCATTTTCATGC 1740
Db 1612 CCTGGAAAACCATCACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTCATTTTCATGC 1671
Qy 1741 TGAAGAGCAGCAGAGAGAGAAAATTCATGAGGAAAAAGGAGCAACTGGCATTTGCAGCTGGC 1800
Db 1672 TGAAGAGCAGCAGAGAGAGAAAATTCATGAGGAAAAAGGAGCAACTGGCATTTGCAGCTGGC 1731
Qy 1801 AGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAGGAGGAGCTCTTGATGGAGAT 1860
Db 1732 AGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAGGAGGAGCTCTTGATGGAGAT 1791
Qy 1861 GCAGAGTCTGTGTTGGGGCGAGAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1920
Db 1792 GCAGAGTCTGTGTTGGGGCGAGAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1851
Qy 1921 AGGAGCTGAGGACAGGAGCTGGCGCAACAGCGGAATATTCGATTTCCTTCTGCCCCAA 1980
Db 1852 AGGAGCTGAGGACAGGAGCTGGCGCAACAGCGGAATATTCGATTTCCTTCTGCCCCAA 1911
Qy 1981 GTGTGAGAGGTTCTCCCTGACATAGACAGTTTACAGATTCACGTGATGATTCATCAT 2040
Db 1912 GTGTGAGAGGTTCTCCCTGACATAGACAGTTTACAGATTCACGTGATGATTCATCAT 1971
Qy 2041 TTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2077
Db 1972 TTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2008
```

RESULT 8

AK055403
LOCUS
DEFINITION
Accession
Version
Keywords
Source
Organism

AK055403
Homo sapiens cDNA FLJ30841 fis, clone FEBRA2002508, highly similar to Homo sapiens FIP2 alternatively translated mRNA.
AK055403
AK055403.1 GI:16550123
oligo capping; fis (full insert sequence).
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Db	316	CAGAGAAA	CAGAAGGAAGAACGCCAGTTTTTTTTCAGATACAGAGCAAAAGACAAAGAGC	375
Qy	597	GTCTAATG	CCCTTGAGTCATGAGATGAGAAATTCGAAGCAAGAGCTTGGAAAACATAAAG	656
Db	376	GTCTAATG	CCCTTGAGTCATGAGATGAGAAATTCGAAGCAAGAGCTTGGAAAACATAAAG	435
Qy	657	GGAAATC	AGAAAGGTTCATCTGAGGACCCCACTGATGATCTCCAGGCTTCCAGGGCCGAAG	716
Db	436	GGAAATC	AGAAAGGTTCATCTGAGGACCCCACTGATGATCTCCAGGCTTCCAGGGCCGAAG	495
Qy	717	CGGAGCAG	AAAAAGGACAGCTCAGGACCCAGGTGTGAGGCTTCAAGCAGAGAGGCGAG	776
Db	496	CGGAGCAG	AAAAAGGACAGCTCAGGACCCAGGTGTGAGGCTTCAAGCAGAGAGGCGAG	555
Qy	777	ACCTGTGG	CGCATCGTGTCTGACCTGACCTCAAGCTGAACCTCCAGCGGCTCTCTCAGAG	836
Db	556	ACCTGTGG	CGCATCGTGTCTGACCTGACCTCAAGCTGAACCTCCAGCGGCTCTCTCAGAG	615
Qy	837	ATTCCCTT	GTGTGAAATTTAGGATGGCTGAAGGGAAGCAGAAAGGGTCAGTAAAGAAATCA	896
Db	616	ATTCCCTT	GTGTGAAATTTAGGATGGCTGAAGGGAAGCAGAAAGGGTCAGTAAAGAAATCA	675
Qy	897	AGCATAGT	CCTGTGGGCCACGAGAACAGTCTCCACTGGCAGCGGATTTGTCTAAATATAGGA	956
Db	676	AGCATAGT	CCTGTGGGCCACGAGAACAGTCTCCACTGGCAGCGGATTTGTCTAAATATAGGA	735
Qy	957	GCAGATCT	GCAGATGGGCGCCAGAAATTTACTTCCGAAATGAGGAGTTAACTGTGAGCCAGC	1016
Db	736	GCAGATCT	GCAGATGGGCGCCAGAAATTTACTTCCGAAATGAGGAGTTAACTGTGAGCCAGC	795
Qy	1017	TCCTGTCT	GTGCTCCTAAGGGAAGGGAATCAGAAAGTGGAGACATTTGAAGTTGCACTCAAGG	1076
Db	796	TCCTGTCT	GTGCTCCTAAGGGAAGGGAATCAGAAAGTGGAGACATTTGAAGTTGCACTCAAGG	855
Qy	1077	AGGCCAA	AGAAAGAGTTTCAGATTTTGAAGAAAAACAAGTAATCGTTCTGAGATTGAAA	1136
Db	856	AGGCCAA	AGAAAGAGTTTCAGATTTTGAAGAAAAACAAGTAATCGTTCTGAGATTGAAA	915
Qy	1137	CCGACAG	CAGAGGGGACAGAGAAAGAGAAATCATGAGAGAAAGGCCGCGAGACATGTTG	1196
Db	916	CCGACAG	CAGAGGGGACAGAGAAAGAGAAATCATGAGAGAAAGGCCGCGAGACATGTTG	975
Qy	1197	GAAGCA	AGTGGAGCACTGAACTCCAGGTGACATCTCTGTTTAAAGGACCTTCAAGAGG	1256
Db	976	GAAGCA	AGTGGAGCACTGAACTCCAGGTGACATCTCTGTTTAAAGGACCTTCAAGAGG	1035
Qy	1257	CTCATACA	AAACTCAGCGAAGCTGAGCTAATGAAGAAAGACATTTCAAGAAAAGTGTGAGG	1316
Db	1036	CTCATACA	AAACTCAGCGAAGCTGAGCTAATGAAGAAAGACATTTCAAGAAAAGTGTGAGG	1095
Qy	1317	CCCTTGA	AGGAAAAATTCGCAATTCATCCAGGTTGAATGAAACCAAGAGAGCTTGTGTT	1376
Db	1096	CCCTTGA	AGGAAAAATTCGCAATTCATCCAGGTTGAATGAAACCAAGAGAGCTTGTGTT	1155
Qy	1377	ATACTACA	AAAAAGTTAGAGCTACAAGTGAAGACATGCTATCAGAAATCAAAATGGAAC	1436
Db	1156	ATACTACA	AAAAAGTTAGAGCTACAAGTGAAGACATGCTATCAGAAATCAAAATGGAAC	1215
Qy	1437	AGGCTAAA	ACAGAGGATGAAAGTCCAAATTAATCTGTCTACAGATGACACACAAAGC	1496
Db	1216	AGGCTAAA	ACAGAGGATGAAAGTCCAAATTAATCTGTCTACAGATGACACACAAAGC	1275
Qy	1497	TTCTTCA	AGACATAAATGCAATTTGAGGAACTTGAAGGAACTTCAAGAAAAAGAGTCAG	1556
Db	1276	TTCTTCA	AGACATAAATGCAATTTGAGGAACTTGAAGGAACTTCAAGAAAAAGAGTCAG	1335
Qy	1557	AAAAAGT	GGACAGGCGAGTGTGAGGAACTGAGTGAAGAACTTGGAACTTGGCAGAGAAGG	1616
Db	1336	AAAAAGT	GGACAGGCGAGTGTGAGGAACTGAGTGAAGAACTTGGAACTTGGCAGAGAAGG	1395
Qy	1617	CTCTGG	CTTCCAAACAGCTGCAAAATGAAATGAAGCAAAACATTTGCCAGCAGGAG	1676
Db	1396	CTCTGG	CTTCCAAACAGCTGCAAAATGAAATGAAGCAAAACATTTGCCAGCAGGAG	1455
Qy	1677	AGGACCT	CGAAACCATGACCATCTCAGGGCTCAGATGGAAGTACTTACTGTTCTGATTTTC	1736
Db	1456	AGGACCT	CGAAACCATGACCATCTCAGGGCTCAGATGGAAGTACTTACTGTTCTGATTTTC	1515
Qy	1737	ATGCTGAA	AGAGCAGCAGAGAGAAAATTCATGAGAAAAGGCAACTGGCATTCGACG	1796
Db	1516	ATGCTGAA	AGAGCAGCAGCAGAGAGAAAATTCATGAGAAAAGGCAACTGGCATTCGACG	1575
Qy	1797	TGGCAGT	TTCTGCTGAAAGAGAAATGCTTTTCCAAAGACGAGCGAGGCACTCTTCTGATGG	1856
Db	1576	TGGCAGT	TTCTGCTGAAAGAGAAATGCTTTTCCAAAGACGAGCGAGGCACTCTTCTGATGG	1635
Qy	1857	AGATCAG	AGATCGTCTCATGCGGCGAGAAACAAAGTGAATCTTCAACAGAGGCTTACCTTGTTC	1916
Db	1636	AGATCAG	AGATCGTCTCATGCGGCGAGAAACAAAGTGAATCTTCAACAGAGGCTTACCTTGTTC	1695
Qy	1917	AAAGAGG	AGCTGAGCAGAGGACTGGCGCAACAGCGGAATATTCGATTTCATTCCTGCGC	1976
Db	1696	AAAGAGG	AGCTGAGCAGAGGACTGGCGCAACAGCGGAATATTCGATTTCATTCCTGCGC	1755
Qy	1977	CCAAAGT	GTGAGAGGTTCTGCTGACATAGACACGTTACAGATTTCACGTGATGGATTGCA	2036
Db	1756	CCAAAGT	GTGAGAGGTTCTGCTGACATAGACACGTTACAGATTTCACGTGATGGATTGCA	1815
Qy	2037	TCATTTAA	GTGTTGATGATACCTCCCAAACTTGTGTT	2077
Db	1816	TCATTTAA	GTGTTGATGATACCTCCCAAACTTGTGTT	1856
LOCUS	BC013876 Homo sapiens optineurin, mRNA (cdna clone MGC:10510 IMAGE:3831267), complete cds.			
DEFINITION	BC013876 Homo sapiens optineurin, mRNA (cdna clone MGC:10510 IMAGE:3831267), complete cds.			
ACCESSION	BC013876			
VERSION	BC013876.2 GI:33878316			
KEYWORDS	MGC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 2318) Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klusner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Loguillano, N.A., McEwan, P.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
JOURNAL	12477932			
PUBMED	2 (bases 1 to 2318)			
REFERENCE	Strausberg, R. Direct Submission Submitted (07-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
PUBMED	12477932			
REFERENCE	2 (bases 1 to 2318)			
AUTHORS	Strausberg, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (07-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:15530202.
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC/DCTP/DTMP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadani@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 14 Row: f Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 20149571.

FEATURES

source

1. .2318
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:10510 IMAGE:3831267"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH_MGC_20"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. .2318
/gene="OPTN"
/notes="synonyms: HYPL, TFIIIA-INTP, NRP, HIP7, FIP2, GLCIE"
/db_xref="GeneID:10133"
/db_xref="MIM:602432"
146..1879
/gene="OPTN"
/codon_start=1
/product="optineurin"
/protein_id="AAH13876.1"
/db_xref="GI:15530203"
/db_xref="GeneID:10133"
/db_xref="MIM:602432"
/translations="MSHQPLSLTEKEDSPSSTGNPPHLAHPNLDTPPELLOQ
KELLNHLQKAMKLNQMRFEELSAWTKQKBERQFEIQSKEAKERLMAISH
ENKLEELKGLKRSERSEPTDSDRLPRAEQRKQDLQTVVRLQAEKADLIGI
VSELQKLNSSGSDSFEIRMAEGAEQSVKEIKHSPGPTRTVSTGALSRYRS
ADGAKNYFHEELTVSOLLCLPEGNQKVERLEVALKEKERSVDEPKTSRSELET
QTGSTEKENDKEGPETVGSEVEALNLQVSLFKELQEAHTKLSAEELMKRLQK
QALERNSAIPSELNKEQLVYTNKLELQVESMLSEIKWQAKTEDEKSKLTVLQMT
HNKLQEHNAKLTIELTRKSEKVDRAVLKELSEKLELAERKALSKQLQMDKQF
IAQBEDLETMTILRAQMEVYCSDFHAERAAAREKIHEEKELALQALAVLLKENDAFED
GQSLMEMQSRHGARTSDSDQAYLVQRAEDRWQRNPIHSCPKGCVGLPDD
TLQIHVMDCLII"

gene

CDS

312 AAGCCATGAAGCTAAATAATAACAAGCCATGAAGGGAGATTGTAGGAGCTTTTCGGCCTGGA 371
537 CAGAGAAACAGAGGAGAGAACGCCAGTCTTTTGTAGATACAGAGCAAAAGAGCAAAAGAGC 596
372 CAGAGAAACAGAGGAGAGAACGCCAGTCTTTTGTAGATACAGAGCAAAAGAGCAAAAGAGC 431
597 GTCTAATGGCTTGAAGTCATGAGAAATGAGAAATTTGAAGGAGAGCTTTGGAAGAACTAAAG 656
432 GTCTAATGGCTTGAAGTCATGAGAAATGAGAAATTTGAAGGAGAGCTTTGGAAGAACTAAAG 491
657 GGAATTCAGAAAGGTCTATCTGAGACCCCACTGATGATCTCAGGCTTCCAGAGGCCAAG 716
492 GGAATTCAGAAAGGTCTATCTGAGACCCCACTGATGATCTCAGGCTTCCAGAGGCCAAG 551
717 CGGAGCAGGAGGAGGACCACTCAGGACCCAGGTGGTGGCTACAGCAGAGAGAGGAGCAG 776
552 CGGAGCAGGAGGAGGACCACTCAGGACCCAGGTGGTGGCTACAGCAGAGAGAGGAGCAG 611
777 ACCTGTTGGCCTCGTGTCTGAACTGAGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAG 836
612 ACCTGTTGGCCTCGTGTCTGAACTGAGCTCAAGCTCAAGCTCAAGCTCCTCAGAAG 671
837 ATTCTCTTTGTTGAAATTAGGATGGCTGAGGAGAGAGGAGGAGGAGTCACTAAAGAAATCA 896
672 ATTCTCTTTGTTGAAATTAGGATGGCTGAGGAGAGAGGAGGAGTCACTAAAGAAATCA 731
897 AGCATAGTCTGGGCCCCAGAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGA 956
732 AGCATAGTCTGGGCCCCAGAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGA 791
957 GCAGATCTGCAGATGGGCCCCAAGAAATTAATCTTCCAAATAGAGGAGTTAACTGTGAGCCAGC 1016
792 GCAGATCTGCAGATGGGCCCCAAGAAATTAATCTTCCAAATAGAGGAGTTAACTGTGAGCCAGC 851
1017 TCCTGCTGTGCTTAAGGGAGAGGAAATCAGAAAGTGGAGAGACTTGAAGTTGCACCTCAAGG 1076
852 TCCTGCTGTGCTTAAGGGAGAGGAAATCAGAAAGTGGAGAGACTTGAAGTTGCACCTCAAGG 911
1077 AGSCCAAGAGAGAGTTTCAGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1136
912 AGSCCAAGAGAGAGTTTCAGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 971
1137 CCAGA CAGAGGGGAGCAG 1196
972 CCAGA CAGAGGGGAGCAG 1031
1197 GAAGCAAGTGGAG 1256
1032 GAAGCAAGTGGAG 1091
1257 CTCATACAAACTCAGCAG 1316
1092 CTCATACAAACTCAGCAG 1151
1317 CCCTTGAAG 1376
1152 CCCTTGAAG 1211
1377 ATACTAACAAAGTTAGAGCTACAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1436
1212 ATACTAACAAAGTTAGAGCTACAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1271
1437 AGGCTAAACACAG 1496
1272 AGGCTAAACACAG 1331
1497 TTCTTTCAG 1556
1332 TTCTTTCAG 1391
1557 AAAAAAGTGGAG 1616
1392 AAAAAAGTGGAG 1451

ORIGIN

Query Match 85.7%; Score 1779.4; DB 5; Length 2318;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1780; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 297 CAGGAATCTTGCATCAACCTCTCAGCTGCTCCTCACTGAAAGAGGAGGAGCAGCC 356
DB 132 CAGGAATCTTGCATCAACCTCTCAGCTGCTCCTCACTGAAAGAGGAGGAGCAGCC 191
QY 357 CCAGTGAAGAGCAG 416
DB 192 CCAGTGAAGAGCAG 251
QY 417 CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCCTGACCAAGAGAGAGAGAGAGAGAGAGAG 476
DB 252 CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCCTGACCGAGAGAGAGAGAGAGAGAGAG 311
QY 477 AAGCCATGAAGCTAAATAATAACAAGCCATGAAGGGAGATTGTAGGAGCTTTTCGGCCTGGA 536

Qy	721	GCAGGAAAGCACCAGCTCAGGACCCAGGTCGTGAGGCTACAAGCAGAAAGCGACGCT	780
Db	651	GCAGGAAAGCACCAGCTCAGGACCCAGGTCGTGAGGCTACAAGCAGAAAGCGACGCT	710
Qy	781	GTTGGGCATCGTGTCTGAACTGCAGCTCAAGCTGAACTCCAGCGCTCCTCAGAAGATTC	840
Db	711	GTTGGGCATCGTGTCTGAACTGCAGCTCAAACTGAACTCCAGCGCTCCTCAGAAGATTC	770
Qy	841	CTTTGTTGAAATTAGATGGCTGAAGGAGAAAGACGAGGGTCAGTAAAGAAATCAAGCA	900
Db	771	CTTTGTTGAAATTAGATGGCTGAAGGAGAAAGACGAGGGTCAGTAAAGAAATCAAGCA	830
Qy	901	TAGTCTGGGCCCCACGAGAACAGTCTCCACTGGCACCGCAATTGCTAAATATAGGACG	960
Db	831	TAGTCTGGGCCCCACAGAACAGTCTCATTTGGCAC-----GAGCAG	872
Qy	961	ATCTGCAGATGGGCCAAGAAATTACTTGCAAATGAGGAGTTAACTGTGAGCCAGCTCCT	1020
Db	873	ATCTGCAGAGGGGCCAAGAAATTACTTGGAAATGAGGAGTTAACTGTGAGCCAGCTCCT	932
Qy	1021	GCTGTGCTTAAGGAAAGGGAATCAGAGGTGGAGACTTGAAGTTGCACTCAAGAGAGGC	1080
Db	933	GCTGTGCTTAAGAGAAAGGGAATCAGAGGTGGAGACTTGAATTTGCACTCAAGAGAGGC	992
Qy	1081	CAAAAGAAAGAGTTTCAGATTTTGAAGAAGAAACAAGTAATCGTTCTCGAGATTGAAACCCA	1140
Db	993	CAAAAGAAAGAGTTTCAGATTTTGAAGAAGAAACAAGTAATCGTCTCGAGATTGAAACCCA	1052
Qy	1141	GACAGAGGGGACACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTTGAAG	1200
Db	1053	GACAGAGGGGACACAGAGAAAGAGAAAGAGAGAAAGGCCCGGAGACTGTTTGAAG	1112
Qy	1201	CGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTGAAGAGCTTCAAGAGGCTCA	1260
Db	1113	CGAAGTGAAGCACTGAACCTTCCAGGTGACATCTCTGTTTGAAGAGCTTCAAGAGGCTCA	1172
Qy	1261	TACAAAACCTCAGCGAAGCTGAGCTAAATCAAGAGAGACTTCAAGAAAGTGTCTAGGCGCT	1320
Db	1173	TACAAAACCTCAGTGAAGCGAGCTTAATGAAGAGAGACTTCAAGAAAGTGTCTAGGCGCT	1232
Qy	1321	TGAAGGAAAAAATTCGCAATTCACATCAGAGTTGAAATGAAAGCAAGAGCTGTTTATAC	1380
Db	1233	TGAAGGAAAAAATTCGCAATTCACATCAGAGTTGAAATGAAAGCAAGAGCTGTTTATAC	1292
Qy	1381	TACAAAAAGTTAGAGCTACAGTGGGAAAGCATGCTATCGAAATCAAAATGGAACAGGC	1440
Db	1293	TACAAAAAGTTAGAGCTACAGTGGGAAAGCATGCTCTCGAAATCAAAATGGAACAGGC	1352
Qy	1441	TAAAAACAGAGGATCAAAAGTCCAAATTAACGTGCTACAGATGACACACACAGAGCTTCT	1500
Db	1353	TAAAAACGAGGATGAAAAGTCCAAATTTAGCCATGTTACGTTTGACACACACAGAGCTTCT	1412
Qy	1501	TCAAGAACATAATATGCAATTTGAAAAAAATTTGAGGAACTAACAGAAAAAGAGTCAGAAAA	1560
Db	1413	TCAGGAACATAATGCAATTTGAAAAAAATTTGAGGAGCTAACAGAAAAAGAGTCAGAAAA	1472
Qy	1561	AGTGGACAGGCGAGTGTGAGGAACTGAGTGAAGAAACTGGAACTGGCAGAGAGGCTCT	1620
Db	1473	AGTGGACAGGCGAGTGTGAGGAACTGAGTGAAGAAACTGGAACTGGCAGAGAGGCTCT	1532
Qy	1621	GGCTTCCAAAACAGCTGCAATTTGAATGAAATGAAACCAATTCGCAAGCAGGAAGAGGA	1680
Db	1533	GGCTTCCAAAACAGCTGCAATTTGAATGAAATGAAACCAATTTCCCAAGCAGGAAGAGGA	1592
Qy	1681	CCTGGAAACCATGACCATCTCTCAGGCGCTCAGATGGAAAGTTTACTGTGTTCTGATTTTCAATGC	1740
Db	1593	CCTGGAAACCATGACCGCTCTCTCAGGCGCTCAGATGGAAAGTTTACTGTGTTCTGATTTTCAATGC	1652
Qy	1741	TGAAGAGCAGCGAGAGAGAAAAATTCATAGGAAAAAGGAGCAACTGGCAATTGCAAGCTGGC	1800
Db	1653	TGAAGAGCAGCGAGAGAAAAATTCATAGGAAAAAGGAGCAACTGGCAATTGCAAGCTGGC	1712

Qy	1801	AGTTCTCCTGAAAGAGAAATGATGCTTTTCGAAGACGAGGAGGAGGAGTCCCTTGATGGAGAT	1861
Db	1713	AGTTTTGCTGAAAGAGAAATGATGCTTTTCGAAGATGGAGGAGGAGTCCCTTGATGGAGAT	1772
Qy	1861	GCAGAGTCGTCATGCGGCGAGAAACAAGTGACTCTGACCCAGCAGGCTTTACCTCTGTTCAAAG	1920
Db	1773	GCAGAGTCGTCATGCGGCGAGAAACAAGTGACCCCGACCCAGAGGCTTACCTTGTTCAAAG	1832
Qy	1921	AGGAGCTGAGGACAGGGAAGCTGGCGGCAACAGCGGAATATTTCCGATTCAATTCCTGCCCCAA	1980
Db	1833	AGGAACTGAGGACAGGGAAGCTGGCAGCAACAAGCGGAATATTTCCGATTCAATTCCTGCCCCAA	1892
Qy	1981	GTGTGGAGAGGTTCTCCCTGACATAGACAGTTACAGATTACGTCATGATGGAATTCGCATCAT	2040
Db	1993	GTGTGGAGAGGTTCTGCCTGACATGACAGCTACAGATTACGTCATGATGGAATTCGCATCAT	1952
Qy	2041	TTAAGTGTTCGATGATATCACCTCCGCCAAAACCTGTTGGT	2077
Db	1953	TTAGGTGTTCGATGATGTCACCTCCGCCAAAACCTGTTGGT	1989
RESULT 12			
LOCUS	CQ771153	1734 bp	DNA
DEFINITION	Sequence 1 from Patent EPL388590.	linear	PAT 04-MAR-2004
ACCESSION	CQ771153		
VERSION	CQ771153.1	GI:45125286	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Kouchi, Y., Masago, A. and Takahata, T.		
TITLE	Gene assay method for predicting glaucoma onset risk		
JOURNAL	Patent: EP 1388590-A 1 11-FEB-2004;		
	Sysmex Corporation (JP)		
FEATURES			
source	Location/Qualifiers		
	1..1734		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
	Query Match	83.3%;	Score 1730.8; DB 2; Length 1734;
	Best Local Similarity	99.9%;	Pred. No. 0;
	Matches 1732;	Conservative 0;	Mismatches 2; Indels 0; Gaps 0;
Qy	311	ATGTCCCATCAACCTCTCAGCTGCCTCAGTGAAGGAGGACAGCCCGAGTGAAGCACA	370
Db	1	ATGTCCCATCAACCTCTCAGCTGCCTCAGTGAAGGAGGACAGCCCGAGTGAAGCACA	60
Qy	371	GGAATGGACCCCGCCACCTGGCCCAACCCAAACCTCGACACGTTTACCCCGGAGGAGCTG	430
Db	61	GGAATGGACCCCGCCACCTGGCCCAACCCAAACCTCGACACGTTTACCCCGGAGGAGCTG	120
Qy	431	CTCAGCAGATGAAGAGCTCTGACCAAGAACCAACAGCTGAAGAGCCATGAAGCTA	490
Db	121	CTCAGCAGATGAAGAGCTCTGACCGGAACCAACAGCTGAAGAGCCATGAAGCTA	180
Qy	491	AATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTCCGCCCTGGACAGAGAACAGAG	550
Db	181	AATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTCCGCCCTGGACAGAGAACAGAG	240
Qy	551	GAAGAACGCCAGTTTTTTTCAGATACAGAGCAAGAGAGGCTCTAATGGCCTTG	610
Db	241	GAAGAACGCCAGTTTTTTTCAGATACAGAGCAAGAGAGGCTCTAATGGCCTTG	300
Qy	611	AGTCATGAGAAATGAGAAATTGAAGGAAGAGCTTGGAAAATCTAAAGGGAAATCAGAAAG	670
Db	301	AGTCATGAGAAATGAGAAATTGAAGGAAGAGCTTGGAAAATCTAAAGGGAAATCAGAAAG	160

/note="similar to human optineurin (OPTN), mRNA,
NM_001008211"
/codon start=1
/product="hypothetical protein"
/protein_id="BA833067.1"
/db_xref="GI.13358908"
/translation="MSHQPUSCLTEKGDSPSESTGNPPHLAHPNLDFTPEELLQOM
KEULTENHOJKEAMKUNNOAMKGRFELSABTKQKEEROFFETQKEAKERLMALSH
ENKJKEELKJLKGKERSSEDDTDRSLPRAEAEQKDLQRTQVRLQAEKADLJGI
VSLQLKLSNGSSSEDSFVIRMAEGEAGSVKEIKHSPGPTVTSIGTSRSAGAKN
YLHEELTVSOLLCLREGNOKVERLEIALKEKRVSDPEKASNSRSIETOTEGST
EKENEKGPETVGSVEALNLQVTSIFKELOEAPTKLSEALMKRLQKQALBERK
NSATPELNEKQBLVINKLELQVESMLSEIKOEQAKTEDEKSKLAMLQDTHNKLLQ
ENHAKLTIELFRKSEKVDRAVLELSEKLEBAKALASKQLQNDENKQTTAKQEE
DLFTMTVLKAEVYCSDFHAEARAAREKIHKEEKEQLALQALVLLKENDAFEDGGRSL
MEMQSRHGARTSDPDQOAYLVQRGTEDRDWQQQRNIP IHSCKPCGEVLPDIDTLQIHV
MDCII"

ORIGIN

Query Match		80.5%; Score 1672.6; DB 5; Length 2681;
Best Local Similarity		91.7%; Pred. No. 0;
Matches 1838; Conservative		0; Mismatches 79; Indels 88; Gaps 3;
QY	73	CGSCAGGTGTGGCTTTTGATAGCTGGTGGTCCCACTTCTCGCCCTTGGATCAGCGGTACGC 132
DB	366	CGGAGGTGTGGCTTTTGACACCTGGTGGTGCCACTTCTCGGTCTTGGATGAGCTGTACGC 425
QY	133	CTCTGTAAACCCCACTTCTCACTTTGAAACACAGCTGCCTGGTTTCAGCATTAATAGAT 192
DB	426	CTCTGTAAACCCAGCTTCTCACTAT-AAAACAGCTGCCTGGTTTCAGCTTAATAGAT 484
QY	193	TAGTCAGTGACAGGCTGTGTGCTGAGTCCGCACATAGAGAATCAAAAATGTCAAAA 252
DB	485	TAGTCAGTGACAGGCTGTGTGTCGAGTCCGCACATA----- 522
QY	253	TGTAACGTGAGAGAAAGTGGGCACTTTTGAGTGACTTTTCCACAGGAACCTTCTGCAAT 312
DB	523	-----GAACTTCTGCAAT 535
QY	313	GTCCCATCAACCTTCTCAGCTGCCTCACTGAAAGAGGAGCAGCCAGTGAAAGCACAGG 372
DB	536	GTCCCATCAACCTTCTCAGCTGCCTCACTGAAAGGGGAGCAGCCAGTGAAGCACAGG 595
QY	373	AAATGGACCCCCCACTGTGCCCCACCAACCTTGGACACAGCTTTACCCCGGAGAGCTGCT 432
DB	596	AAATGGACCCCCCACTGTGCCCCACCAACCTTGGACACATTCACCCAGAGAGCTGCT 655
QY	433	GCAGCAGATGAAGAGCTTCTGACCAAGAACCCACAGCTGAAGAGGCCATGAAGCTAAA 492
DB	656	GCAGCAGATGAAGAGCTTCTGACCCGAGAACCCACAGCTGAAGAGGCCATGAAGCTAAA 715
QY	493	TAATCAAGCCATCAAGGGGAGATTTCAGGAGCTTTCGGCTTGACAGAGAAACAGAGGA 552
DB	716	TAACCAAGCCATCAAGGGGCGGTTTGAAGGAGCTTTCGGCTTGACAGAGAAACAGAGGA 775
QY	553	AGAACCCAGTTTTTTTCAGATACAGAGCAAGAGCAAGAACAGCGCTCTAATGGCTTGAG 612
DB	776	AGAACCCAGTTTTTTTTCAGATACAGAGCAAGAGCAAGAACAGCGCTCTAATGGCTTGAG 835
QY	613	TCATGAGATGAAGAAATTGAAGGAGAGCTTTGAAAACTAAAAAGGAAATTCAGAAAGGTC 672
DB	836	TCACGAGATGAAGAAATTGAAGGAGAGCTTTGAAAACTGAAAACTGAAAAAGGAAATTCAGAAAGGTC 895
QY	673	ATCTGAGGACCCCACTGATGACTTCCAGGCTTCCAGGCGCGAGCCGAGCGAGAAAGGA 732
DB	896	ATCTGAGGACCCCACTGATGACTTCCAGGCTTCCAGGCGCGAGCGAGCGAGAAAGGA 955
QY	733	CCAGCTCAGGACCCAGGTGTGAGGCTTACAAGCAGAGAGGCGAGACCTGTTGGGCATCGT 792
DB	956	CCAGCTCAGGACCCAGGTGACGAGGCTACAGCAGAGAGGCGAGACCTGTTGGGCATGCT 1015
QY	793	GTCTGAACCTGACGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAAGATTTCCTTTGTGAAAT 852

DB	1016	GTCTGAACCTGCAGCTCAAACTGAACCTCAGCGGCTCTCGGAAGACTCCTTTGTGAAAT 1075
QY	853	TAGGATGGCTGAAGGAGAGACAGAGGGTCAGTAAAGAGAAATCAAGCATAGTCTCGGGCC 912
DB	1076	TAGGATGGCTGAAGGAGAGACAGAGGGTCAGTAAAGAGAAATCAAGCATAGTCTCGGGCC 1135
QY	913	CACGAGAAACAGTCTCCATCTGGCAACGGCATTTGTCTAAATATATAGGAGCAGATCTCCAGATGG 972
DB	1136	CACAGAAACAGTCTCCATCTGGCAC-----GAGCAGATCTCGAGAGG 1177
QY	973	GGCAGAAATTAAGTTCGAACATCAGAGAGTTAACTGTGAGCCAGCTCTGCTGTGCCTAAG 1032
DB	1178	GGCAAGAAATTAAGTTCGAACATCAGAGAGTTAACTGTGAGCCAGCTCTGCTGTGCCTAAG 1237
QY	1033	GGAAAGGGAATCAGAAAGGTGAGAGACTTCAAGTTGACACTCAAGAGGAGCCAAAGAGAGT 1092
DB	1238	AGAAAGGGAATCAGAAAGGTGAGAGACTTCAAGTTGACACTCAAGAGGAGCCAAAGAGAGT 1297
QY	1093	TTCAGATTTTGAAGAGAAACCAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGAG 1152
DB	1298	TTTCAGATTTTGAAGAGAAACCAAGTAATCGCTCTGAGATTGAAACCCAGACAGAGGGAG 1357
QY	1153	CACAGAGAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGAAGC 1212
DB	1358	CACAGAGAAAGAGAAACCGAGAGAAAGGTCAGAGACTGTTGGAAGCGAAGTGAAGC 1417
QY	1213	ACTGAACTCCAGGTGACATCTCTGTTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAG 1272
DB	1418	ATTGAACTTTCAGGTGACATCTCTGTTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAG 1477
QY	1273	CGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAAGTGTCAAGGCCCTTTGAAAGGAAAA 1332
DB	1478	TGAAGCCGAGCTTAATGAAGAGAGACTTCAAGAAAAAGTGTCAAGGCCCTTTGAAAGGAAAA 1537
QY	1333	TTTCGCAATTCATCAGAGTTGAAAGAGAGAGCTTGTATATACTAAACAAAAGTT 1392
DB	1538	TTTCGCAATTCATCAGAGTTGAAAGAGAGAGCTTGTATATACTAAACAAAAGTT 1597
QY	1393	AGAGCTTACAAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACACAGGA 1452
DB	1598	AGAGCTTACAAAGTGAAGAGCATGCTCTCGAAATCAAAATGGAACAGGCTTAAACACAGGA 1657
QY	1453	TGAAAAGTCCAAATTAACCTGTGTACAGATGACACACAAACAGCTTCTTCAAGAAACATAA 1512
DB	1658	TGAAAAGTCCAAATTAACCTGTGTACAGATGACACACACACACAGCTTCTTCAAGAAACATAA 1717
QY	1513	TAATGCAATTTGAAAAACAAATTTGAGAGAACTAAACAGAAAGAGTCAGAAAAAGTGACAGGGC 1572
DB	1718	TCATGCAATTTGAAAAACAAATTTGAGAGAACTAAACAGAAAGAGTCAGAAAAAGTGACAGGGC 1777
QY	1573	AGTGTCTGAAGGAACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCTGGCTTCCAAACA 1632
DB	1778	AGTGTCTGAAGGAACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCTGGCTTCCAAACA 1837
QY	1633	GCTGCAATTTGGAATGAAGCAACCACTTCCAGAGCAGGAGAGGACCTGGAACCAAT 1692
DB	1838	GCTGCAATTTGGAATGAAGCAACCACTTCCAGAGCAGGAGAGGACCTGGAACCAAT 1897
QY	1693	GACCAATTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCACTGCTGAAAGAGCAGC 1752
DB	1898	GACCGTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCACTGCTGAAAGAGCAGC 1957
QY	1753	GAGAGAGAAATTCATGAGGAAAGGAGCACTGGCAATTCAGCTGGCAGTCTGCTGAA 1812
DB	1958	AGAGAGAAATTCATGAGGAAAGGAGCACTGGCAATTCAGCTGGCAGTCTGCTGAA 2017
QY	1813	AGAGAAATTCATGAGGAGGAGGAGGAGTCTTTCATGAGGAGATGAGAGTCTGCTCA 1872
DB	2018	AGAGAAATTCATGAGGAGGAGGAGGAGTCTTTCATGAGGAGATGAGAGTCTGCTCA 2077
QY	1873	TGGGGCGAGAAACAGTACTCTGAACAGAGGCTTACCTTGTTCAAAGAGGAGCTGAGGA 1932
DB	2078	TGGGGCGAGAAACAGTACTCTGAACAGAGGCTTACCTTGTTCAAAGAGGAGTCTGAGGA 2137

QY	1495	GCTTCTTCAAGAACATATAATGTCATGTAATAAATGAGAACTAAGAGAAAGAGTC	1554
Db	1304		
QY	1555	AGAAAAGTGGACAGCGCAGTCTGAAGCACTGAGTGAAGAACTGGAACTGGCAAGAA	1614
Db	1364		
QY	1615	GGCTCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAAAACCATGGCAAGCAGGA	1674
Db	1424	GGCCCTGGCTTCCAAACAGCTCCAAATGGATGGATGAAGCAGACCATTTGGCAAGCAGGA	1483
QY	1675	AGAGGACCTGGAAAACCATGACCATCTCAGGGCTCAGATGGAAGATTTACTGTTCTGATTT	1734
Db	1484	GAAGGACCTGGAGACCATGGCTGTCTCAGGGCTCAGATGGAGGTTTATTGTTCTGATTT	1543
QY	1735	TCATGCTGAAGAGCAGCAGAGAGAGAAATTCATGAGGAAAAGGAGCAACTGGCATGCA	1794
Db	1544	TCATGCTGAAGAGCAGCAGAGAGAGAAATTCATGAGGAAAAGGAGCAACTGGCATGCA	1603
QY	1795	GCTGGCAGTCTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAG---GCAGGCAGTCCTT	1851
Db	1604	GCTGGCAGTCTTCTGCTGAAAGAAACAAATGCTTTTGAAGATGGAGGCAGCAGCAGACCTT	1663
QY	1852	GATGGAGATGCGAGATGCTCATGGGGCGAGAAACAAGTGAAGTCTTGACACAGAGCTTACCT	1911
Db	1664	GATGGAATGCGAGCGCTCATGGGGCGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	1723
QY	1912	TGTTCAAGAGGAGCTGAGCAGAGGAGCTGGGGCAGCAG---CGGAATATTCGAGTTCA	1968
Db	1724	CGTTCAAGAGGAGGAGGAGGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	1783
QY	1969	TTCTCTGCCCCAAGTGGAGAGGTTCTGCTGACATAGACACGTTACAGATTACAGTGAT	2028
Db	1784	TTCTCTGCCCCAAGTGGAGAGGTTCTGCTGACATAGACACGTTACAGATTACAGTGAT	1843
QY	2029	GGATTGCATCATTTAAGTGTGATGATCATCCTTCCCAAACTGTTGGT	2077
Db	1844	GGACTGTATCATTTAAGTGTGACCTTTTACCTCCCCCAAGCTGTTGGT	1892
RESULT 17			
AF513722			
Sus scrofa optineurin (OPTN) mRNA, complete cds.			
AF513722.1 GI:31324102			
Sus scrofa (pig)			
Sus scrofa			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus			
1 (bases 1 to 1799)			
Obazawa,M., Mashima,Y., Sanuki,N., Noda,S., Kudoh,J., Shimizu,N., Oguchi,Y., Tanaka,Y. and Iwata,T.			
Analysis of porcine optineurin and myocilin expression in trabecular meshwork cells and astrocytes from optic nerve head			
Invest. Ophthalmol. Vis. Sci. 45 (8), 2652-2659 (2004)			
15277488			
2 (bases 1 to 1799)			
Sanuki,N., Obazawa,M., Mashima,Y., Tanaka,Y. and Iwata,T.			
Direct Submission			
Submitted (20-MAY-2002) National Institute of Sensory Organs, National Tokyo Medical Center, 2-5-1 Higashiagoka, Meguro, Tokyo 152-8902, Japan			
Location/Qualifiers			
1..1799			
/organism="Sus scrofa"			
/mol type="mRNA"			
/db xref="taxon:9823"			
1..1799			
/gene="OPTN"			
gene			

CDS	1..1725	/gene="OPTN"	
	/codon_start=1		
	/product="optineurin"		
	/protein_id="AAP47176.1"		
	/db_xref="GI:31324103"		
	/translation="MSHQPLSCLTEKGDSPTETTGNGPPTLAHPNLDFTPHLLQOMRELLIENHOLKEAMKLNQAMKRFELUSATEKQKEERLPFTOSKAKERUTALSL ENELKOBLSGLKGTGTERFEDLTGDPVPAKAEQVEQLKTQVQAEKADLIGI VSEQLKNSGSPSDSEFIRMAEGEADAAKKEIKTSPGPIRTSDITSKABETRN YLEPELTVSOLLCLCRGNOKVERLEIALKEAKERILDFEKAKDRSTETQTEBHK ROEKEEEKSPETVGSEVMNLQVTTLFKELQEAHTLSEAEKMKRLQEKQALERK NSATPELNEQELLNNKLELQVETSMRSIKMEQAKTEEEKSLTTLQHPNRLQ VYNNAKLTIELKRRESEKDKVLQELNGLEMAEKALASQOLQMDENKQIIAQEK DLETMVLAQMEVYSDPFAERAEREKHEEKQLALQLALLKDDNAFEFGAGRSQ LMEQSRHGARAADAOQAFLVQGAEDRNMLQQQQNIPIHSCPKCEVLPIIDTL IHTVDCII"		
ORIGIN			
	Query Match	62.2%;	Score 1291;
	Best Local Similarity	84.6%;	Pred. No. 0;
	Matches 1495;	Conservative	0; Mismatches 245; Indels 27; Gaps 3;
QY	311	ATGTCCCATCAACCTCTCAGCTGCTCACTGAAAAGGAGGACAGCCCAAGTGAAGACACA	370
Db	1	ATGTCCCATCAACCTCTCAGCTGCTGACTGAAAAGGGGAGCAGCCCAAGTGAAGACACA	60
QY	371	GGAAATGGACCCCCCACTGCGCCCAACCCAAACCTGGACACGTTTTACCCCGGAGGAGCTG	430
Db	61	GGAAATGGACCCCCCACTGCGCTCACCACCAACCTTGACACGTTTCAACCCACATGAACCTG	120
QY	431	CTGCAGCAGATGAAGAGCTCTGACCAAGAAACCAAGCTGAAAGAACCCATGAAGCTA	490
Db	121	CTGCAGCAGATGAGAGAGCTTCTAATTCGAGAAACCATCAGCTGAAAGAACCCATGAAGCTA	180
QY	491	ATAATCAAGCATGAAAGGAGATTTGAGGAGCTTTGGCGCTGGACAGAGAAACAGAAAG	550
Db	181	ATAATCAAGCTATGAAAGGGCGATTTGAGGAGCTTTGAGCGCTTGAAGAGAGCAGAAAG	240
QY	551	GAAAGAACGCCAGTTTTTTGAGATACAGAGCAAGCAAGCAAGAGAGCGCTCTAATGGCGCTTG	610
Db	241	GAAAGAACGCCCTTTTTTTGAGACCCAGAGCAAGCAAGAGCGCTTACCGGCTCTG	300
QY	611	AGTCATGAGAAATGGAAGAGCTTTGGAAGAACTTGGAAAACTAAAAGGGAATCAGAAAGG	670
Db	301	AGTCTTGAATAAGAAAACTGAAGCAAGAACTTGGAAAACTAAAAGGGAATCAGAAAGG	360
QY	671	TCATCTGAGGACCCCACTGATGACTCCAGGCTTCCCGAGGCGCAAGCGGAGGAGGAAAG	730
Db	361	TCATTTGAGGACCTCACTGGGAGCCCGAGGCTCCCAAGCGGCGCAAGCGGAGGAGGAAAG	420
QY	731	GACCACTCAGGACCCAGGTGTGTGAGGCTTACAAGCAGAGAGGAGGAGCAGCTTTGGGCATC	790
Db	421	GACAGCTGAAGACCCAGGTGTGACGCTTCAAGCTGAAAAGCGGATCTGCTGGGCATC	480
QY	791	GTGTCTGAAGTCTGAGCTCAAGCTGAATCTCCAGCGGCTCTCTCAGAGATTCCTTTGTGAA	850
Db	481	GTGTCTGAATTCAGCTCAAGCTGAATCTCAGGTGGCGCCCTCTGAAGACTCTTTGTGAA	540
QY	851	ATTAGAGTGGCTGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	910
Db	541	ATCAGAGTGGCTGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	600
QY	911	CCCAAGAGAACTGCTCCCACTGGCAGCGCATTTGTCTAAATATATAGGAGCAGATCTGCAGAT	970
Db	601	CCCAATAGAACTGATTCATTTGACAC-----GAGCAATCTGCAGAA	642
QY	971	GGGGCCAGAAATTAATTCGAAATGAGGAGTTAACTGTGAGCGCAGCTCTCTGTGTGCTTA	1030
Db	643	GGTACCAGGAATTAATTTGGAATTTGAGGAATTAATCTGTGAGCCAGCTCTCTGTGTGCTTA	702
QY	1031	AGGGAAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCGCAAGAAAGA	1090

Db 703 AGGGAAGAAACCAAGAGGTGGAGAGACTTGAATCGCCCTCAAGGAAGCCAAAGAAAGA 762
Qy 1091 GTTTTCAGATTTTGAAGAAACCAAGTAATCGTTCTGAGATTGAAACCCACAGACAGAGGG 1150
Db 763 ATTTTAGATTTTGAAGAAAGCCAGGATCGTTCTGAGACTGAGACCCACAGACAGAGAG 822
Qy 1151 AGCAGAGAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAGCGGAAGTGGA 1210
Db 823 CACAAGAACACAGAGAAAGAGAGAGAGAAAGCCAGAAAACCTGTTGGAAGTGAAGTGGA 882
Qy 1211 GCATCGAACCTTCAGGTCGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATACAAAATC 1270
Db 883 ATGTTAAACCTTCAGGTCGACAACTGTTTAAAGAGCTTCAGGAGGCTCACACGAATC 942
Qy 1271 AGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTCAAGGCCCTTGAAGAGAA 1330
Db 943 AGTGAAGCTGAGCTCATGAAGAGAGACTTCAAGAAAATGTCAAGGACCTTGAAGAGAA 1002
Qy 1331 AATTCTGCAATTCATCAGATTGAATGAAGAAAGAGAGCTGTTGTTTATATACAAAAG 1390
Db 1003 AATTCTGCAACCCCATCAGAGCTGAATGAAGAAAGCAGAGCTTCTTTATATAACAAAAG 1062
Qy 1391 TTGAGCTACAGTGAAGAGCATCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAG 1450
Db 1063 TTGAGCTCAAGTGAAGAGCATGAGATCAGAAATTAATAATGGAAGCAAGCCAAAACAG 1122
Qy 1451 GATGAAAGTCAAAATTAATGCTGTACAGATGACACACAAAGCTTCTTCAAGAACAT 1510
Db 1123 GAGGAAAGTCAAAATTAATGCTGTACAGATGACACACAAAGCTTCTTCAAGAACAT 1182
Qy 1511 AATATGATGAAAGCAATTTGAGAACTAACAAGAAAAGAGTGCAGAAAAGTGGACAGG 1570
Db 1183 AATAATGCATGAAACAAATTTGAGAACTGAAAAGAGAGAGTCTGAAAAGTGGATAAG 1242
Qy 1571 GCAGTCTGAGGAACTGAGTGAAGAACTGGAACCTGCAGAGAGGCTCTGGCTTCCAAA 1630
Db 1243 GTGGTCTGAGGAACTGAGTGAAGAACTGGAAGCTGGAAGCTGGAAGAGGCTTGGCTTCAAG 1302
Qy 1631 CAGCTGCAAAATGGAATGAAGAACTGGAAGCTGGAAGCTGGAAGCTGGAAGCTGGAAGCT 1690
Db 1303 CAGCTCCAAATGGAATGAGATGAAGCAAGCACTTCCCAAGCAGAGAGGAGCTGGAAGCT 1362
Qy 1691 ATGACCATCTCAGGCTCAGATGGAAGTGTATCTGTTCTGATTTTCATGCTGGAAGAGCA 1750
Db 1363 ATGCTCTGTTCTCAGGCTCAGATGGAAGTGTATCTGTTCTGACTTTTCATGCTGGAAGAGCA 1422
Qy 1751 GCGAGAGAGAAATTCATGAGGAAGAGCAAGCACTGCAATTCAGCTGGCAGTTCGCTG 1810
Db 1423 GCAAGAGAGAGATTCATGAGGAAGAGGAGCACTGCAATTCAGCTGGCAGTTCGCTG 1482
Qy 1811 AAAGAGAAATGATGCTTTGGAAGAGGAG--GCAGGAGCTCTTGTATGGAAGATGCAGAGT 1867
Db 1483 AAAGAGCAATGCTTTTGAAGAGGAGGAGCCAGAGGCAATCTTGTATGGAAGATGCAGAGC 1542
Qy 1868 CGTCATGGGCGAGAAACAAGTACTCTGACAGAGGCTTACCTTGTTCAGAGAGGAGCT 1927
Db 1543 CGTCATGGGCGAGAGCAAGTATGCTGACAGAGGCTTCTTGTTCAGAGAGGAGCT 1602
Qy 1928 GAGGACAGGAGCTGGCGGCAAG-----CGGAAATTCGATTCATTCCTGCGCCCAAG 1981
Db 1603 GAGGATGAAGCTGGCTGCAACAGCAACAACAGAAATTCCTCAATTCATTCCTGCGCCCAAG 1662
Qy 1982 TGTGGAGAGGTTCTGCTGACATAGACACAGTTACAGATTCACGCTGATGATTCATTCATT 2041
Db 1663 TGTGGAGAGGTTCTGCTGACATAGATACACTGATTCACGCTGATTCATTCATTTCATTCATT 1722
Qy 2042 TAAGTGTGATGATACCTCCCAAA 2068
Db 1723 TAAGTGTGATGATTTTACTTCCCAAA 1749

RESULT 18
BC061185

LOCUS BC061185 2443 bp mRNA linear ROD 30-JUN-2004
DEFINITION Mus musculus optineurin, mRNA (cDNA clone MGC:74325 IMAGE:30314331), complete cds.
ACCESSION BC061185
VERSION BC061185.1 GI:38511794
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE 1. (bases 1 to 2443)
Srausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L.H., Shennen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, M., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Loguellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smaluk, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2. (bases 1 to 2443)
Strausberg, R.
Direct Submission
Submitted (03-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaphs@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ahgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 53 Row: p Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction, Similarity but not identity to protein.
FEATURES
source Location/Qualifiers
1. .2443
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:74325 IMAGE:30314331"
/tissue_type="Kidney, mouse"
/clone_lib="NIH_MGC_176"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
1. .2443
gene

```
/gene="Optn"  
/note="synonyms: NRP, FIP2, HYPL"  
/db_xref="GeneID:71648"  
/db_xref="MGI:1918898"  
44.._1798  
/gene="Optn"  
/codon_start=1  
/product="optineurin"  
/protein_id="AAH61185.1"  
/db_xref="GI:38511795"  
/db_xref="GeneID:71648"  
/db_xref="MGI:1918898"  
/translation="MSHQPLSLCTEKGDSPCTPGNSPNWVHPSLDFTPELLLOQM  
KELLVHOLKEAMKLNINQAMKRFEEISAWTEKQERLLFEMOSKEVEKRLKALH  
ENERLKEELGKEKSEKPLEDLTGYPYRALBEEVEKLTQVEVHEHLKIQVRL  
RAEKADLLGIVSELOKLNSSGSSDSFVIRMTGEGTEGAMKMKNCPTPTPTDPI  
LSEKTDARSACFEELTVSOLLCLREGNQKVERLVALREAKERISDFEKKANGHS  
STEKOTAPRADREKDKGOESVSEVETLSIOVTSLPKELQEAHTKLSEAEILMKRLQ  
EKQALERNKSNATBELNEKQELVYSNKKLELOVESMRSEIKHEQAETEEKSRLATL  
QATHKLQEDHNKALKTIELTKQAKVDKMLLQELSEKLESLASOALASKOLQNDDEM  
KQTLAKQEDHETMAVLRQMEVYCSDFHAERAAREKTHEEKELQALQALAILLKEND  
IEEGSRQSLMEMQCRHGARTSDSDQTYLFGAEDRSWQHGOQPRSIPIHSCPKCG  
EVLPTDITLIHWMDCII"
```

ORIGIN

Query Match 53.1%; Score 1102.4; DB 6; Length 2443;
Best Local Similarity 77.9%; Pred. No. 7.6e-282;
Matches 1420; Conservative 0; Mismatches 341; Indels 63; Gaps 5;

QY	268	AGTGGCAACTTTTGAGTGACTTTTCCACAGAACTTCTGCAATGTCTCCCATCAACCTCT	327
DB	1	AGTGACCTAAGAGGAGAGGATCCCTGTGGCCGAGCTTACCATGTCCCATCAACCTCT	60
QY	328	CAGCTCCCTCACTGAAAGAGGAGACAGCCCGAGTGAAGCAGAGAAATGGACCCGCCA	387
DB	61	GAGCTCCCTGACTGAGAGGGGACAGCCCTTGTGAGACCCAGGAANTGGACCTCCAA	120
QY	388	CCTGGCCCAACCCAAACCTTGACACAGTTTACCCCGAGGAGCTGCTGCAGCAGATGAAGA	447
DB	121	TATGGTTTCAACCCAGCCTGGACACATTCACCCCTGAGGAGCTGCTGCAGCAAAATGAAGA	180
QY	448	GCTCCTGACCAAGAACCCAGCTGAAGAAGCCATGAAGCTAAATAATCAAGCCATGAA	507
DB	181	ACTCCTGGTTGAGAACCAACAGCTGAAGAAGCCATGAAGCTTAATAATCAAGCTATGAA	240
QY	508	AGGGAGATTTGAGGAGCTTTGGGCTGGACAGAGAAACAGAGAAAGAACCGCAGTTTTT	567
DB	241	AGGGCGATTTGAGGAGCTGTCCGCTGGACAGAGAGCAGAGAGAGAGCGCCTGTGTT	300
QY	568	TGAGATACAGACAAAGACAAAGAGCGGTCTAATGGCGCTTGAGTCATGAGAAATGAGAA	627
DB	301	TGAGATGCAAGCAAGAGAGGTTAAGGAGCGCCTTAAGGCGCTGACTCATGAAATGAGAG	360
QY	628	ATTGAGGAGAGCTTGGAAACTTAAAGGGAATCAGAAAGCTCATCTGAGGACCCAC	687
DB	361	GCTGAAGGAGAGCTTGGAAATTCAGAGAGAAATCAGAAAGCCATGGAAGACCTCAC	420
QY	688	TGATGACTCCAGGCTTCCACAG-----GGCCGAAGC	717
DB	421	AGTGGCTACAGTATCCACAGACCTTGGAGGAGGAAGTGGAGAACCTGAAGACCCAGGT	480
QY	718	GGAGCAGGAAAGAGACAGCTCAGGACCCAGGTGGTAGGCTCAAGAGAGAGAGGAGAG	777
DB	481	GGAGCAGGAAAGTGGAGCATCTGAAGATCCAGGTGATCGCGCTTCGGGCTGAAAGGCAGA	540
QY	778	CCTGTTGGGCATCGTCTGAACCTGCAGCTCAAGCTGAACCTCAGCGGCTCCTCACAAGA	837
DB	541	CCTGCTGGGCATCGTCTCAGACTGCAGCTCAAACTCAACTCCCGCGGCTTCCTCGGAAGA	600
QY	838	TTCTCTTTGTAATAGGATGGCTGAAGGAGAGCAGAGGGTCAAGTAAAGAAATCAA	897
DB	601	CTCTCTTGTGATCAGGATGATCCGAAGAGAGAGACTGGAAGGGCCATGAAGGATGAA	660

QY	898	GCATAGTCTCGGCCACAGAGAACAGTCTCCACTGGCAGCGCATTTGTTCTAAATATAGGAG	957
DB	661	GAACTGCCCTACACCCACAGAAACAGACCCCATCAGC-----TTGAG	702
QY	958	CAGATCTCGAGATGGGCCCAAGAAATTACTTCGAAACATGAGGAGTTAACTGTGAGCCAGCT	1017
DB	703	CAACTGTACAGAGGATGCCAGGAGTTGTGCGGAGTTTGAAGAACTGACTGTGAGCCAGCT	762
QY	1018	CCTGCTGTGCTTAAGGGAAGGAATCAGAAAGTGTGAGAGACTTGAAGTTGCACTCAAGGA	1077
DB	763	TCTGCTTTGCCCTAAGGGAAGGAAACCAAAAGGTGTGAGAGACTTGAAGTGCCTCAGAGA	822
QY	1078	GGCCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAATCGTTCTGAGATTGAAC	1137
DB	823	AGCCAAAGAAAGAAATTCAGATTTTGAAGAAAGAAACAAATGGCCATTTCTTCTACTGAGA	882
QY	1138	CCAGACAGAGGGGAGCACAGAGAAAGAAATGATGAAGAAAGGCCCGGAGACTGTTGG	1197
DB	883	CGACACAGGAGGAGAGCAGACAGAGAGAGAAAG---GGACAAAGGCCCAAGAGAGTGTGG	939
QY	1198	AAGCGAAGTGAAGACCTGAACTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGC	1257
DB	940	AAGCGAAGTGGAAACACTGAGCATTTCAAGTGCACCTCTCTGTTTAAAGGAGCTTCAAGAGGC	999
QY	1258	TCATACAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTGAGGC	1317
DB	1000	ACACACAAACTCAGTGAGGCTGAGCTGATGAAGAGAGACTTCAAGAAAGTGTGAGGC	1059
QY	1318	CCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAAAAGCAAGAGCTTGTGTTA	1377
DB	1060	TCTGAGAGGAGAACTCTGCAACACCATCAGAGCTGAATGAAAAAGCAAGAGCTGTTTA	1119
QY	1378	TACTAACAAAAAGTTAGAGCTTCAAGTGAAGAGATGCTATCAGAAATCAAAATGGAACA	1437
DB	1120	CAGTAACAAAGAGTTAGAGCTGAGGTGAGAGCATGCGCTCCGAAATCAAGTAGAGCA	1179
QY	1438	GGCTAAACAGAGGATGAAGTCCAAATTAACCTGTGTACAGATGACACACAAAGCT	1497
DB	1180	GGCCAAAGCAGAGGAGGAGAGTCCAGGTTAGCCACTCTGCGAGGCACTCACAACAGCT	1239
QY	1498	TCTTCAAGAACATAATAATGCAATTGAAACCAATGAGGAACTTAACAAGAAAAAGTFCAG	1557
DB	1240	CCTTCAAGAACATAATAAGGCACTGAAACCAATGGAAGAACTAACCAAGCAACAGGCAGA	1299
QY	1558	AAAAAGTGAACAGGAGCTGCTGAAGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGC	1617
DB	1300	AAAGGTGGACAAAGATGTTGCTGCAGGAGCTCAGCGAGAAAGCTGGAGCTGGCAGAGAGGC	1359
QY	1618	TCTGCTTCCAAACAGCTGCAATGGATGAATGAAGCAAAACCATTTGCCAAGCAGGAAGA	1677
DB	1360	TCTGGCATCCAAACAGCTCCAGTGGATGAGATGAAGCAGACGCTCGCTTAGCAGGAGGA	1419
QY	1678	GGACCTTGGAAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCA	1737
DB	1420	AGACCTTGGAGACCATGGCCGCTCCTCAGGCTCAGATGGAAGTGTACTGCTCAGATTTTCA	1479
QY	1738	TGCTGAAGAGAGCAGCAGAGAGAAATTCATGAGGAAAGGAGGACAACTGGCATTCAGCT	1797
DB	1480	CGCTGAGAGAGCAGCAGAGAGAAATTCATGAAGAAAGGAGAGCAGCTGGCCTTTCAGCT	1539
QY	1798	GGCAGTTCTGCTGAAGAGAAATGATGCTTTTCCGAAGACGAGGC---AGCAGTCTTGTAT	1854
DB	1540	CGCGATTTTGTCTGAAGAGAACATGAATGAAGGAGGAGGAGGAGTACAGTCTTCTGAT	1599
QY	1855	GGAGATGCAAGAGTCTGATGGGGCCGAGAAACAAGTGAAGTCTGACACAGCAGCTTACCTGT	1914
DB	1600	GGAAATGCAAGTCCGACACGGGGCAAGAACCAAGTGAAGTCTGACACAGCAGACTTACCTGT	1659
QY	1915	TCAAAGAGAGCTGAGGACAG-----GGACTGGCGGCAACACAGCGGAATATTCGAT	1965
DB	1660	TCAAAGAGAGCGGAGGACAGGAGCTGGCAGCACGGGCGAGCAGCTCCGAGTATTCGAT	1719
QY	1966	TCATTCCTCCCAAGTGTGAGAGGTTCTGCTGACATAGACACGTTTACAGATTCACGCT	2025

Db 1720 TCACTCTGCCCCAAGTCGGAGAGGTCCTGCGGACATGCACACGCTTCAGATCCATGT 1779
Qy 2026 GATGGATTGCATCATTTTAAGTGTT 2049
Db 1780 GATGGACTGCATCATTTTGAAGTGTT 1803

RESULT 19
AY071834
LOCUS Mus musculus optineurin (Optn) mRNA linear ROD 06-JUL-2002
ACCESSION AY071834
VERSION AY071834.1 GI:21700320
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2023)
Rezaie,T. and Sarfarazi,M.
TITLE Cloning and Characterization of Mouse Optineurin
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2023)
Rezaie,T. and Sarfarazi,M.
AUTHORS Direct Submission
TITLE Submitted (26-DEC-2001) Molecular Ophthalmic Genetics Laboratory,
JOURNAL Surgical Research Center, Department of Surgery, University of
Connecticut Health Center, 263 Farmington Ave, Farmington, CT
06030-111, USA

FEATURES
source
1. .2023
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129"
/db_xref="taxon:10090"
/chromosome="2"
/map="within 6.6 cR and between D2Mit76 and D2Mit354"
/cell_type="dermal fibroblast"
gene
1. .2023
/gene="Optn"
5'UTR
1. .150
/gene="Optn"
151..1905
/gene="Optn"
/note="OPTN"
/codon_start=1
/product="optineurin"
/protein_id="AAL61853.1"
/db_xref="GI:21700321"
translations="MSHOPLSCLTEKGDSPCTPGNGPSNMVHPSLDTPPELLOQM
KELLVENHQLKEAMKLNQAMKGRFELSNATEKQKBEKLLFEMQSKVEKRLKALTH
ENRLEKEELFKFEKSEKPLEDITGGYRPRALREBEVKLKTQVEVHLKIQVRL
RAEKDLGTVSLQKLNKSGGSEDSFVIEIRMTGETEGAMKEMKNCPTPTDPTIS
LSNCTDARSACFEELTVSOLLCLREGNOKVERLVALREAKERISDFEKKANGHS
SKEOTRADREDKDGQSVSEVETLSIOVSLFKELQEAHTKLSAEELMKKELQ
EKQALERNKATSPSELNEKQELVYNKKLELOVESNRSEIKNEQAKTBEKSRLATL
QATHKLLQHNKALKTIETLTKQAKVKMDLQJLQELSEKLELAELQALASKQLQMDM
KQTLAKOEDLETMVLRQMEVYCSDFHAERAAEKIHEEKQALQALAILKENND
IEEGSRQSLMENQCRHGARTSDSDQTVLFRGABDRSNWQHQPRSIPIHSCPKCG
EVLPDIDTLQIHVMDCII"
1906..1956
/gene="Optn"

ORIGIN
Query Match 53.0%; Score 1101.8; DB 6; Length 2023;
Best Local Similarity 77.5%; Pred. No. 1.2e-281;
Matches 1439; Conservative 0; Mismatches 352; Indels 66; Gaps 6;
Qy 263 GAGAAAGTGGGCAACTTTTGGATGACTTTTTCACAGGAACTTCTGCAATGTCCCATCAA 322
Db 103 GAAGCAGTGACCTTAAGAGGAGGATCCCTGTGGCGGACCTGTTCACATGTCCCATCAA 162

Qy 323 CCTCTCAGTGTGCTCACTGAAAAAGGAGGACAGCCCGAGTCAAGACACAGGAAATGAGCCC 382
Db 163 CCTCTGAGCTGCTGCTGACTGAGAAAGGGGACAGCCCTTGTGAGACCCCGGAAATGAGACC 222
Qy 383 CCCCACTGTGCCCCACCCAAACCTGGACACGTTTATCCCCGGAGGAGTGTCTGCAGCAGATG 442
Db 223 TCCAATATGTTTCACCCCAGCTGGACACATTCACCCCTCAGGAGTGTCTGCAGCAATG 282
Qy 443 AAGAGCTCTCTGACCAACAGACACCAAGCTGAAAGAGCCATGAGCTTAAATTAATCAAGCC 502
Db 283 AAGAACTCTCTGTTGAGAACCCACAGCTGAAAGAGCCATGAGCTTAAATTAATCAAGCT 342
Qy 503 ATGAAAGGAGATTTTGGAGAGCTTTTGGCTGTCACACAGAGAAACAGAGGAGAAACGCGCAG 562
Db 343 ATGAAAGGCGATTTGAGGAGCTGTCCGCTGACACAGAGAAAGCAGAGGAGAGCGCGCTG 402
Qy 563 TTTTGTGAGATACAGAGCAAAAGAAAGAGCGTCTAATGCGCTTTAGTCTGATGAGAAT 622
Db 403 TTGTTTGGATGCAAAAGCAAAAGAGTTTAAAGAGCGCTTAAAGGCCCTGACTCATGAAAT 462
Qy 623 GAGAAATTGAAAGAGAGCTTTGAAACTTAAAGGGGAAATCAGAAAGGTCTATCTGAGGAC 682
Db 463 GAGAGGCTGAAGAGAGAGCTTTGGAAATTCAGAGAGAAATCAGAAAGCCATTTGGAAGAC 522
Qy 683 CCCACTGATGACTCCAGGCTTTCCAG-----GGCC 712
Db 523 CTCACAGGTGGCTACAGGTATCCAGAGCTTTGGAGGAGGAGTGGAGAAAGCTGAAGACC 582
Qy 713 GAACCGGAGCAGGAAAGGACCAAGCTCAGGACCCAGGTGGTGGGTGTACAGAGAGAG 772
Db 583 CAGGTGGAGCAGGAAAGTGGAGCATCTGAAAGTCCAGGTGATGCGCTTCGGGCTGAAAG 642
Qy 773 GCAGACTGTGTTGGGCATCGTCTGAACTGCAGCTCAAGCTGAACCTCCAGCGGCTCTCTCA 832
Db 643 GCAGACTGTGTTGGGCATCGTCTCAGAACTGCAGCTCAAACTCAACTCCGCGGCTCTCTCG 702
Qy 833 GAAGATTCTCTTTCTCAAAATTAGGATGGCTGAAAGGAGAGACAGAAAGGTCAGTAAAGAA 892
Db 703 GAAGACTCTCTTCTTGATCAGGATGACCGAAGGAGAGACTGAGGGGCAATGAAGGAG 762
Qy 893 ATCAAGCATAGTCTCTGGGCCACAGAGAACAGTCTCCACTGGCACCGGCACTTGTCTAAATAT 952
Db 763 ATGAAGAACTGCGCTTACACCCCAAGAACAGACCCCATCAGC----- 804
Qy 953 AGGAGCAGATCTGCAGATGGGGCCAAAGAAATTACTTCGAAACATGAGAGATTAACTGTGAGC 1012
Db 805 TTGAGCAACTGTACAGAGGATGCCAGGAGTTGTGCGGAGTTTGAAGAACTGACTGTGAGC 864
Qy 1013 CAGCTCTCTGTGCTTAAGGGAAGGAAATCAGAAAGTGGAGAGACTTGAAGTTGCACCTC 1072
Db 865 CAGCTTCTGCTTTGCTTAAGGGAAGGAAACCAAAAGTGGAGAGACTTGAAGTGCGCCCTC 924
Qy 1073 AAGGAGGCCAAAGAAAGACTTTTCAAGATTTTGAAGAAAGAAACCAAGTAAATCGTCTGAGATT 1132
Db 925 AGAAGAGCCAAAGAAAGATTTTCAAGATTTTGAAGAAAGAAATGCGCCATTTCTTCTACT 984
Qy 1133 GAAACCCACACAGAGGGGAGCAGAGAAAGAGATGATGAAGAGAAAGCCCGGAGACT 1192
Db 985 GAGAAGCAGACACGAGGAGAGCAGACAGAGAGA---AGGAGGACAAAGGCCAAGAGAGT 1041
Qy 1193 GTTGAAGCGAAGTGGAAAGCACTGACCTCCAGGTCACATCTCTGTTTAAAGGAGCTTCAA 1252
Db 1042 GTTGAAGCGAAGTGGAAAGCACTGACATTTCAAGTACCTCTCTGTTTAAAGGAGCTTCAA 1101
Qy 1253 GAGGCTCATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAAGTGT 1312
Db 1102 GAGGCACACACAAAACCTCAGTGGGCTGAGCTGATGAAGAGAGACTTCAAGAAAAAGTGT 1161
Qy 1313 CAGGCCCTTGAAGGAAAAATTTTCAATTCATCAGAGTTGAATGAAAGCAAGAGCTT 1372
Db 1162 CAGGCTCTGAGGAGGAAAGAACTCTGCAACCATCAGAGCTGAATGAAAGCAAGAGCTC 1221

[illegible]

Qy 898 GCATAGTCTCTGGGCCACGAGAACAGTCTCACTGGCACGCAATGTCTAAATATAGAG 957
Db 601 GAAACAGCGCTGGACCCACGAGACAGACTCCATCAGCATGG-----G 642
Qy 958 CAGATCTCAGATGGGGCCAAAGATTAATCTTCAACATGAGAGTTAACTGTAGCCAGCT 1017
Db 643 CAATGTGACAGAGACGCGAGACTTGTGTGGAGTTTGGGGAACCTGACTGTGAGCCAACT 702
Qy 1018 CCTGTGTGCTTAAGGGAAGGGAATCAGAAGGTGAGAGACTTGAAGTTGCACTCAAGGA 1077
Db 703 CTTGCTTTGCTCAGGGGAAGAACCAACAAAGGTGGAGACTCGAGTCGGCTTCAGAGA 762
Qy 1078 GGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACCAAGTAAATCGTTCTGAGATTGAAC 1137
Db 763 AGCCAAAGAAAGAAATTTTCAGATTTTGAAGAAACCAAGCGCAATCTCTGCGATTGAGAC 822
Qy 1138 CCAGACAGAGGGGACGACAGAGAAAGAGAAATGATGAGAGAAAGCCCGGAGACTGTGG 1197
Db 823 CCAGACAGAGGGGACACAAAAAG---AAGAGGAGACAAAGACCCGGAGAGTGTGG 879
Qy 1198 AAGCGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGC 1257
Db 880 AATCGAAGTGAAGACTCTGAACGTTCAAGTGGCTCTCTGTTTAAAGAGCTTCAAGAGGC 939
Qy 1258 TCATACAAAACCTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAGTGTCAAGC 1317
Db 940 GCACACAAAGCTCAGTGAGGCGGAGCTGATGAAGAAGAGACTTCAAGAAAAGTGTCAAGC 999
Qy 1318 CTTTGAAGGAAAATTTCTGCAATTCATCAGAGTTGAATGAAGCAAGAGCTGTGTTA 1377
Db 1000 TCTGGAAGGAAAGAACTCTGCAACCCATCGAGCTGAATGAAAGCAAGAGCTCGTTA 1059
Qy 1378 TACTAACAAAAGTTAGAGCTTACAGTGAAGAGCATGCTACAGAAATCAAAATGGGAACA 1437
Db 1060 CAGTAACAGGAAGTTAGCTGCGAGGTGAGAGCATGCGCTCGGAAATCAAGATGGAGCA 1119
Qy 1438 GGCTAAACAGAGGATGAAAAGTCCAAATTAACCTGTGCTACAGATGACACACAACAAGCT 1497
Db 1120 GGCCACAGACAGAGGAGGAATCCAGGTTAGCCACTCTTCAGGCAACACACAGCAAGCT 1179
Qy 1498 TCTTCAAGAACATTAATGATGTAAGAAACATTTGAGGAACATAACAGAAAAGAGTCAGA 1557
Db 1180 TCTTCAAGAACACATAAAGCTCTGAGAACAAATTAAGAACTAAACCAACAAACAGGCAGA 1239
Qy 1558 AAAAGTGCACAGGCGAGTGTGAGGAACTGAGTGAAACCTGAACTGGCAGAGAGGC 1617
Db 1240 AAAAGTGCACAGTGTGAGGAGCTGCGAGGAGCTCAGCGAAGCTGGAGCTGGCGGAGCAGC 1299
Qy 1618 TCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAACCAATTCGAAGCAGGGAAGA 1677
Db 1300 TCTGGCTTCCAAAGCAGCTCCAGATGGATGAGTGAAGCAGACCATGCCAAGCAGGAGGA 1359
Qy 1678 GGACTGGAACCATGACATCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTCA 1737
Db 1360 GGACTGGAACCATGAGCGCTCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTCA 1419
Qy 1738 TGCTGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGAGGCAACTGGCATTCAGCT 1797
Db 1420 TGCTGAGAGGCGACGAGGAGAGAGATCCATGAGGAAAGAGCAGCTGGCGCTTGCAGCT 1479
Qy 1798 GGCAGTTCTGTGAAAGAGATGATGCTTTTGAAGAGCGGAGGC---AGGCGAGCTCTTGAT 1854
Db 1480 CGCCATTTTGTGAAAGAGAACATGACTTTGAGATGGAGGAGTAGGAGCCCTTGAT 1539
Qy 1855 GGAGATGAGAGTGTCTATGGGGCGAGAACAGTGAATCTGACACAGAGGCTTACCTGT 1914
Db 1540 GGAATGAGTGTGCGGACGCGGGCGAGAACCCAGGCACTCTGACACAGAGGCTTACCTGT 1599
Qy 1915 TCAAGAGGAGCTGAGACAGAGGAGCT-----GGCGGCAACAGCGGGAATATTCGAT 1965
Db 1600 TCAAGAGGAGCGAGGACATGAGCTGGCAGCATGAGGCGAGCAGCGCCCTGAGTATTCCTAT 1659

Qy 1966 TCATTCTCTGCCCAAGTGTGAGAGGTTCTGCTGACATAGACACGCTTACAGATTCA 2022
Db 1660 TCATTCTGCCCCAAGTGTGGGAGGTCCTGCGGACATTTGACACGCTTACAGATA 1716
RESULT 21
AB069907 3577 bp mRNA linear ROD 19-SEP-2003
DEFINITION Rattus norvegicus mRNA for FIP-2 like protein, complete cds.
ACCESSION AB069907
VERSION AB069907.3 GI:34850479
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE
1 Oyama, M.
Identification of genes expressed in rat injured dental pulp
Unpublished
2 (bases 1 to 3577)
Oyama, M. and Myokai, F.
Direct Submission
Submitted (11-AUG-2001) Masataka Oyama, Nagasaki University
Graduate School of Biomedical Sciences, Division of Oral Health
Services Research; 1-7-1 Sakamoto, Nagasaki, Nagasaki 852-8588,
Japan (E-mail: masataka@net.nagasaki-u.ac.jp, Tel: 81-95-849-7663,
Fax: 81-95-849-7665)
On Sep 18, 2003 this sequence version replaced gi:21743244.
COMMENT
FEATURES
Location/Qualifiers
1..3577
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
27..1787
/notes="alternative spliced FIP-2"
/codon_start=1
/product="FIP-2 like protein"
/protein_id="BAB8496.2"
/db_xref="GI:21743245"
/translation="MSHQPLSLCTEKGDSSCTPGNPSNMVHPLNDFTFPBELLQOM
KELLVENHQLKAMLNQAMKGRFELSAWTERQKLEFBIQSEKAEKRLKALSH
ENRLKEELKLEKSERPEDI TGRGFPRTDLEQVETBAGGAGSAGSDPGBA
PSGLRKTLLGHSSOTAAOQLRLLRLRGDRMTEGEAGAKMNRNSAGPTTDS
IIMGKCTEDARTCVFEELTVSOLLCLREGNOKVERLEALREAKERISDFPKANG
HSAETQTGSGTOKBEDKDPESVGLVETLVNQVASFGLQEAHTKUSEAEUMKCR
LOEKCQALRKNSATSELNKEQLVYSNWKLEQVESMRSEIKNQAKTEEEKQLMD
TLOATHDKLLQEHNKALRTIEELTKQQAEBKVDKVLQELSEKLEAEQALASKQLA
EMKOTAKQBEDLETWAVLRQMEVYCSDFHAERAAREKTHEEKEQLALOLAILLKEN
NDFEDGSGRSQSLMEMQCRHGARTSDSDQAYLFFQRGKMSWQHQQPRSIPIHSCPK
CGEVLDPDITLQIHVMDCII"

Query Match 51.0%; Score 1058.8; DB 6; Length 3577;
Best Local Similarity 77.4%; Pred. No. 2.8e-270;
Matches 1417; Conservative 0; Mismatches 337; Indels 76; Gaps 8;
Qy 297 CAGGAACTTCTGCAATGTCCCATCAACCTCTCAGCTGCTCCTCACTGAAAAGAGGACAGCC 356
Db 13 CCGGACCTGTACCATGTCCCATCAACCTCTGAGCTGCTCCTGAGAGGGGACAGCT 72
Qy 357 CCAGTGAAGACAGAGAAATGGACCCCGCCACCTGCGCCACCCCAACCTGGACACGTTTA 416
Db 73 CCTGTGAGACCCGAGAAATGGACCTTCCCAATATGTTTACCCCAACCTGGACACATTC 132
Qy 417 CCGGAGGAGCTGCTGCAGCAGATGAAGAGCTCTCTGACCAAGAACCAACAGCTGAAG 476
Db 133 CTCTGAGGAGCTGCTGCAGCAATGAAGAACTCTCTGTCGAGAACCAACAGCTGAAG 192
Qy 477 AAGCCATGAAGCTTAAATAATCAAGCCATGAAGGAGATTTGAGGAGCTTTCGCCCTGGA 536
Db 193 AAGCCATGAAGCTTAAATAATCAAGCTATGAAGGGCGATTTGAGGAGCTTTCGCCCTGGA 252

ORIGIN

QY	537	CAGAGAAA	CAGAAAGGAAACCGCCAGT	TTTTTTTGGATACAGAGCAAAAGAAAGAGC	596
Db					
QY	253	CAGAGAGG	CAGAAAGGAAAGCCCACT	TTGTTTGAGATACAAAGCAAAGAGGCTAAGGAGC	312
Db					
QY	597	GTCATAATG	CGCTTTGAGTCATGAGAAATGAGAAAT	TGAAGGAAGAGCTTGGAAAACTTAAAG	656
Db					
QY	313	GCCTCAAGG	CCCTTGAGTCAATGAAATGAGAGGCT	GAAAGGAAGAGCTTGGAAAACTTAAAG	372
Db					
QY	657	GGAATATC	AGAAAGAGTCACTCTGAGGACCCCACTGATGACT	TCCAGGCTTCCACGGGCGC---	713
Db					
QY	373	AGAAATCA	AGAAAGGCCATTTGAGACATCAGAGGTAGATG	CGGGTTTTCCAGAACCGACT	432
Db					
QY	714	-----	-----	-AAGCGGAGCAGGAAAGAACCAAGCTCAGG	742
Db					
QY	433	TGGAGCAGG	AAAGTGGGAGCAACTGAAGAGGCGAGGTGGAGCAGG	AAAGTGGAGCATCTGAAG	492
Db					
QY	743	ACCCAGTG	TGGTGAAGCTACAAGC--	-AGAGAAGGCAGACCTGTTGGGCATCGTGTCTGAA	799
Db					
QY	493	ATCCAGTG	GAGGCGCCTTCAGGCTTCAGAAAGCGGACCTTGCT	GGGGCAATTCGTCCTCAGA	552
QY	800	CTGCAGCT	CAAGCTGAATCTCAGCGGCTCCTCAGAAAGATT	CCTTTGTTGAAATTAGGATG	859
Db					
QY	553	CTGCAGCT	CAAGCTCAACTCCGCGGCTCCTCAGAAAGACT	CCTTCTGTGAGA-CAGGATG	611
Db					
QY	860	GCTGAAGG	AGACAGAGGTCAGTAAAGAAATCAAGCATAGT	TCCTGGGCCACAGAGA	919
Db					
QY	612	ACTGAAGG	AGAGCGGAAGGGCAATGAAGGAGATGAGGAA	CAGCGCTGGACCCACAAGG	671
Db					
QY	920	ACAGTCTC	CACTGGCAGCGCATGTTCTAAATATAGGAGCAGAT	CTCCAGATGGGGCCAAAG	979
Db					
QY	672	ACAGACTC	CAATCATATG-----	GCAAAATGTACAGAGGACGCCAGG	713
Db					
QY	980	AATPACTT	CGAAATGAGGAGTTAACTGTGAGCCAGCTCCTGCT	GTGCTTAAGGGAAGG	1039
Db					
QY	714	ACTTGTGT	GAGTTTGAGGAACCTGACTGTGAGCCAACTCCTGCT	TTTGGCTCAGGGHAGA	773
Db					
QY	1040	AATCAGAG	GTGGAGAGCTTGAAGTTGCACCTCAAGGAGGCGCA	AAAGAGAGTTTCAGAT	1099
Db					
QY	774	AACCAAGG	TGGAGAGACTCGAGATCGCGCTCAGAGAAGCC	AAAGAAAGTAATTCAGAT	833
Db					
QY	1100	TTTGAAGA	AAAAACAAGTATCTGTTCTGAGATTGAACCC	CAGACAGAGGGGAGCAGACAG	1159
Db					
QY	834	TTTGAAGA	AAAAACAAGCCCACTCTCGGATTGAACCC	CAGACAGAGGGGAGCAGACAA	893
Db					
QY	1160	AAAGAGAA	TGATGAAGAAAGCCCGGAGACTGTTTGAAGCG	AAAGTGAAGCACTGGAAC	1219
Db					
QY	894	AAAG---	AAAGAGGAGGCAAAAGCCAGAGAGTGTGGAA	ATCGAAGTGAACACTCTGAAC	950
Db					
QY	1220	CTCCAGGT	GACATCTCTGTTTAAAGGAGCTTCAAGAGGCT	CATACAAAACCTCAGCGAAGCT	1279
Db					
QY	951	GTTCAAGT	GGCCTCTCTGTTTAAAGGGCTTCAAGAGGCGC	CACAAAAGCTCAGTGAGGCC	1010
Db					
QY	1280	GAGCTAAT	GAAAGAGAGACTTCAAGAAAAGTGTCAAGGCC	CTTGAAGGAAAAATTCGCA	1339
Db					
QY	1011	GAGCTGAT	GAAAGAGAGACTTCAAGAAAAGTGTCAAGGCT	CTCGAAGGAAAGCAACTCTGCA	1070
Db					
QY	1340	ATTTCCAT	CAGAGTTGAATGAAAGCAGAGCTTGTTTATCT	ATAACAAAAGTTACAGCTA	1399
Db					
QY	1071	ACCCCAT	CGGAGCTGAATGAAAGCAGAGCTCGTTTAC	AGTAACTGGAAGTTAGAGCTG	1130
Db					
QY	1400	CAAGTGGAA	AGCATGCTATCAGAAATCAAAATGGAA	CAGGCTTAAACAGAGGATGAAAG	1459
Db					
QY	1131	CAGGTGAG	GAGCATGCGCTCGGAAATCAAGATGGAGCAG	GCGCCCAAGACAGAGGAGAGAA	1190
Db					
QY	1460	TCCAAATTA	CTGTGCTACAGATGACACAAACAGCTTCTT	CAAGACATATATATGCA	1519
Db					
QY	1191	TCCAGGTT	AGCCACTCTTCTCAGGCAACACAGCAGAC	AGCTCCTTCAGGAACAATAAAGCT	1250
Db					
QY	1520	TTGAAAA	CAATTTCAGGAACCTAAACAAGAAAAAGCT	TCAGAAAAAGTGGACGGGCACTGCTG	1579
Db					
QY	1251	CTGAGAAC	CAATTGAAGAACTTAACAAACCAACAGCGC	AGAAAAAGTGGACAAGGTGAGCTG	1310
Db					

QY	1580	AAGAACTG	AGTGAATAAACTGGAACTGGCAGAGAAAGGCT	CTGGCTTCCAAAACAGCTGCAA	1639	
Db						
QY	1311	CAGGAGCT	CAGCGAAGCTGGAGCTGGCGAGCAGGCT	CTGGCGTCCAAGCAGCTCCAG	1370	
Db						
QY	1640	ATGGATGA	ATAATGAACCAACCATTCGCAAGCAGGAAGAGG	ACCTGGAAACCATGACCATC	1699	
Db						
QY	1371	ATGGATGA	TAATGAACCAACCATTCGCAAGCAGGAAGAGG	ACCTGGAGCCATGGCCGTC	1430	
Db						
QY	1700	CTCAGGGT	CTCAGATGGAAGTTTACTGTTTCTGATTTT	TCATGCTGAAAGAGCAGCAGAGAG	1759	
Db						
QY	1431	CTCAGGGT	CTCAGATGGAAGTTTACTGTTTCTGATTTT	TCATGCTGAGAGGCGCAGCAGGGAG	1490	
Db						
QY	1760	AAAATTC	ATGAGAAAAGAGCAACTGGCATTTGAGCTGGC	AGTTCTCTGTAAGAGAT	1819	
Db						
QY	1491	AGATCCAT	GATGGAAGAGGAGCAGCTGGCTTTGCAAGCT	TCGCGCATTTTCTGTAAGAGAAAC	1550	
Db						
QY	1820	GATGCTTT	TCGAAGACGGAGGC---	AGGCAGTCTTGTATGGAGATGCAGAGT	CGTCTATGGG	1876
Db						
QY	1551	AATGACTT	TGAAGATGGAGGAGTAGGAGTCTTGTATG	GAATGCAGTGCAGGACCGG	1610	
Db						
QY	1877	GGGAGAAC	AAAGTGAATCTGACAGCAGGCTTACCTTGT	TTCAAAAGAGGAGCTGAGGACAGG	1936	
Db						
QY	1611	GGGAGAAC	CCAGGACTCTTGAACAGCAGGCTTACCTGT	TTTCAAAGAGGAGCCAAAGGACATG	1670	
Db						
QY	1937	GAAT-----	GGCGGCAACAGCGGAATATTCGATTCAT	TCTGCCCCCAAGTGTGGA	1987	
Db						
QY	1671	AGCTGGCA	GACATGGGAGCAGCGCCCGGAGTATTC	CCATTCACCTGTCGCCCAAGTGGG	1730	
Db						
QY	1988	GAGGTTCT	CGCTGACATAGACAGCTTACAGATTTAC	GTGATGATTTGCATTTAAAGTG	2047	
Db						
QY	1731	GAGGTCCT	GGCCGGAATGACACAGCTTACATACATG	TATGATGATGATCTATCATCTCAGGG	1790	
Db						
QY	2048	TTGATGAT	ATCACTCCCAAAACTGTTGTT	2077		
Db						
QY	1791	-----	CCCTCCACTCCCAAGCTGTTGTT	1815		
Db						

RESULT 22	BC086976	1897 bp	mRNA	linear	ROD 30-MAR-2005
LOCUS	Rattus norvegicus optineurin, mRNA (cdna clone MGC:93070				
DEFINITION	IMAGE:7111026), complete cds.				
ACCESSION	BC086976				
VERSION	BC086976.1	GI:56268803			
KEYWORDS	MGC.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 1897)				
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Utsid, T.B., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Bosak, S.A., McEwan, P.J., Abramson, R.D., Malek, J.A., Gunaratne, P.H., Richards, S., McKernan, K.J., Huley, S.J., Huley, S.W., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huley, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schermer, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
CONSRMT	Mammalian Gene Collection Program Team				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				

PUBMED 12477932
 REFERENCE 2 (bases 1 to 1897)
 AUTHORS NIH MGC Project
 TITLE Direct Submission
 CONSRM Submitted (02-DRC-2004) National Institutes of Health, Mammalian
 JOURNAL Gene Collection (MGC), Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 182 Row: c Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 37059753.

FEATURES

source
 1..1897
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="MGC:93070 IMAGE:7111026"
 /tissue_type="Testis, rat (Brown Norway)"
 /clone_lib="NIH MGC_237"
 /lab_host="DH10B"
 /notes="Vector: pExpress1"
 1..1897
 /gene="Optn"
 /note="synonym: Fip2"
 /db_xref="GeneID:246294"
 /db_xref="RGD:628886"
 113..1753
 /gene="Optn"
 /codon_start=1
 /product="Optn protein"
 /protein_id="AAH86976.1"
 /db_xref="GI:56268804"
 /db_xref="GeneID:246294"
 /db_xref="RGD:628886"
 /translations="MSHQLPCLITEKGDSSCTPGNGPSNMVHPLDTPPELLOQM
 KELLVNHQLEAMKLNQAMKGRFEELSMTWKEQKEEROLFQISKEAKERLKLQSH
 ENERLKEELKLEKSPEDITGRGPPRTDLEOEVLKQVQEVHVKNQBE
 AEGAMKEMRNSAGPTRDTSLSMGKCTEDATCFEELTVSOLLCLREGNOKVERLE
 IALREAKERLSDEPKKANGHSALETQTEGSTQKEEDKDPESVIGIVETILNVQVSLF
 KELOEAKTULSEALMKRLQERQALERNKSNATPSELNKEQLSVYNSNRKLELVQVSM
 RSEIKMEQATEEKEKRLATLQATHDKLLQEHNKALRTBELTKQQAQKVDKVLQEL
 SEKLELAQALQKQMDQEMKQTIAKQEEDLETMVLRQMEVYVCSDFHAERAREK
 IHEEKQLALAILKENNDFDGGKQSLMEMQCRHGARTSDSDQQAAYLQFQGAEDM
 SWHQGQPRSIPIHSPKCGEVLPLDITLQIHVMDCLII"

ORIGIN

Query Match 48.0%; Score 997.6; DB 6; Length 1897;
 Best Local Similarity 75.7%; Pred. No. 6.3e-254;
 Matches 1355; Conservative 0; Mismatches 319; Indels 116; Gaps 5;
 QY 297 CAGGAATCTTCGATGTCATCAACCTCTAGCTGCCTCCTCTGAAAGAGGACAGCC 356
 DB 99 CCGGACCTGTACCATGTCGCATCACTCTAGCTGCCTCCTCTGAAAGAGGACAGCT 158
 QY 357 CCAGTGAAGACACAGGAAATGGACCCGCCACCTGGCCCAACCCAACTCGACAGCTTTA 416
 DB 159 CTTGTGAGACCCAGGAAATGGACCTTCCAAATATGGTTTACCCCAACCTGGACATTC 218
 QY 417 CCCCAGGAGGAGCTGCTGCAGCAGATGAAGAGCTCCTGACCAAGAACCCACGCTGAAAG 476

Db 219 CTCCTGAGGAGCTGCTGCAGCAAAATGAAGAACTCTCTGGTCCGAGAACCCACGCTGAAAG 278
 QY 477 AAGCCATGAAGCTTAAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTGCGCTTGA 536
 Db 279 AAGCCATGAAGCTTAAATATCAAGCTATGAAGGGGCGATTTGAGGAGCTTTTGCCTTGA 338
 QY 537 CAGAGAAACAGAAAGAGAAACGCCAGCTTTTTCAGATACAGAGCAAAAGAGCAAAAGAGC 596
 Db 339 CAGAGAAACAGAAAGAGAAACGCCAGCTTTTTCAGATACAAAGCAAAAGAGCTTAAGGAGC 398
 QY 597 GTCTAATGGCTTGAAGTCATGAGAAATGAAGAAATTTGAAGAAAGAGCTTTGAAAACTAAAAG 656
 Db 399 GCCTCAAGGCCCTGAGTCATGAATAATGAGAGGCTGAAGGAAGAGCTTTGAAAACTAAAAG 458
 QY 657 GGAATAACAGAAAGAGTCACTCTGAGGAGCCCACTGATGATCCAGGCTTCCAGGCGCCAG 716
 Db 459 AGAATACAGAAAGAGCCATTTGAAGACATCACAGTAGATCGGGGTTTCCAGAAACCCGACT 518
 QY 717 CGGAGCAGGAAAAAGGACCCAGCTCAGGACCCAGGTGGTGAAGCTACAGCAGAGAAAGGCAG 776
 Db 519 TGGAGCAGGAAGTGGAGCACTGAGAGGCGAGGTGGAGCAGGAA----- 562
 QY 777 ACCTGTTGGGCATCGTGTCTGAACCTGAGCTCAAGCTGAACCTCCAGCGGCTCCTCAGAAG 836
 Db 563 ----- 562
 QY 837 ATTCCTTTGTTGAATTAGGATGGCTGAAGGAGAACGAGGCTCAGTAAAGAAATCA 896
 Db 563 -----GTGGAGCATCTGAAGAACGAGGAGGAGGCGGAAGGGCAATGAAGGAGATGA 617
 QY 897 AGCATAGTCTCTGGGCCACAGAGAACAGTCTCCACTGCGACGGCATTTGCTCTAAATATATGA 956
 Db 618 GGAACAGCGCTGGACCCCAAGGACAGACTCCATCAGCATGG----- 659
 QY 957 GCAGATCTGCAGATGGGGCCAAAGAAATTACTTCGAAATGAGGAGTAACTGTGAGCCAGC 1016
 Db 660 GCAAAATGTACAGAGGACGCGCAGGACTTGTGTGAGTTTGAGGAACTGACTGTGAGCCAAC 719
 QY 1017 TCCTGCTGTGCTTAAGGAGGGAATCAGNAGCTGGAGAGACTTGAAGTTGCACTCAAGG 1076
 Db 720 TCCTGCTTTGCTCAGGGAAGGAAACCAAAAGGTGGAGAGACTCGAGATCGCGCTCAGAG 779
 QY 1077 AGSCCAAGAAAGAGTTCAGATTTTGAAGAAAAACAAGTAATCGTTCTGAGATTGAAA 1136
 Db 780 AAGCCAAAGAAAGAAATTTAGATTTTGAAGAAAGAACGACCGCCATTTCTGCCATTTGAGA 839
 QY 1137 CCAGACAGAGGGGAGCAGAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTG 1196
 Db 840 CCCAGACAGAGGGGAGCAGCAAAAG---AAGAGGAGGACAAAGACCCAGAGAGTGTGG 896
 QY 1197 GAGCCAGTGGNAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGG 1256
 Db 897 GAATCGAAGTGGAAACTCTGAAATGTTTCAAGTGGCCCTCTCTGTTTAAAGGAGCTTCAAGAGG 956
 QY 1257 CTCAATAAACTCAGCAGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTCAAGG 1316
 Db 957 CGCACAAAGCTCAGTGAGCGGAGCTGATGAAGAGAGACTTCAAGAAAAAGTGTCAAGG 1016
 QY 1317 CCTTGAAGAGGAAAAATTTCTGCAATTCATCAGAGTTGAATGAAAAAGCAAGAGCTTGT 1376
 Db 1017 CTCTGAAAGGAAGAACTCTGCACCCCATCGAGCTGAATGAAAGCAAGAGCTCGTTT 1076
 QY 1377 ATACTAAACAAAAGTTAGAGCTTACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAAC 1436
 Db 1077 ACAGTAACAGGAAGTTAGAGCTGCAAGTGGAGAGCATGCGCTCGGAAATCAAGATGGAGC 1136
 QY 1437 AGSCTAAAACAGAGGATGAAGTCCAAATTAATCTGCTACAGATGACACACAAAGC 1496
 Db 1137 AGCCCAAGACAGAGGAGGAGAAATCCAGGTTAGCCACTCTTTCAGGCAACACACGACAGC 1196
 QY 1497 TTCTTCAAGAACATAATAATGCAATTGAAAAACAATTCAGGAACTTAACAAGAAAAAGAGTCA 1556

Db 1197 TTCTTCAGGAACACATAAAGCTCTCAGAACAAATTGAAGAACTAACCAACACAGGCAG 1256
Qy 1557 AAAAACTGGACAGGGCAGTGTCTGAAGGAACTGAGTGAAAAAATCTGGAACTGGCAGAGG 1616
Db 1257 AAAAACTGGACAGGGTGCAGTGTGAGGAGCTCAGCGAAGCTGGAGCTGGCGGACAGG 1316
Qy 1617 CTCTGCTTCCAAAAGCTGCAAAATGATGAATGAAGCAACACCATTCACAGCAGGAAG 1676
Db 1317 CTCTGCGTCCAAAGCAGCTCCAGATGATGAATGAAGCAGACCATCGCAAGCAGGAGG 1376
Qy 1677 AGGACCTGGAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTCTCTGATTTTC 1736
Db 1377 AGGACCTGAGACCATGGCGCTCTCAGGGCTCAGATGGAAGTTTACTGTCTCTGATTTTC 1436
Qy 1737 ATGCTGAAAGAGCAGCGAGAGAAATTCATGAGGAAAGAGCAATCTGGCATTCGCAGC 1796
Db 1437 ATGCTGAGAGGCGAGCAAGGAGAGATCCATGAGGAAAGAGCAGCTGGCTTGCAGC 1496
Qy 1797 TGGCAGTCTCTGTAAGAGATGATGCTTTTGAAGACGGAGCGAGCAGTCTGATGG 1856
Db 1497 TCGCCATTTTGTCTGAAAGAGAAACAATGACTTTTGAAGATGGAGCGCAGTCTGATGG 1556
Qy 1857 AGATGACAGTCTCATGCGGCGAGAACAAAGTGAATCTGACAGCAGGCTTACCTTTGTC 1916
Db 1557 AAATGACAGTCCGGCATGGGCGAGAACCAAGGACTCTGACCAAGAGGCTTACCTTTTC 1616
Qy 1917 AAGAGGAGCTGAGGACAGGACT-----GGCGCAACAGCGGAAATATTCGATTC 1967
Db 1617 AAGAGGAGCGGAGACATGAGCTGGCAGATGGCGAGCAGCCCGGATTTCCATTC 1676
Qy 1968 ATTCCTGCCCAAGTGTGAGAGTCTGCTGACATAGACAGCTTACAGATTCAGTGGA 2027
Db 1677 ACTCGTCCCAAGTGTGGGAGGCTCTGCGGACATTGACAGCTTCAGATACATGTGA 1736
Qy 2028 TGGATTGACATTTAAGTGTGATGATCATCACTCCCAAACTGTGTGT 2077
Db 1737 TGGACTGTATCTGAGG-----CCCTCCACTCCCAAGCTGTGTGT 1781

RESULT 23
CQ729566
LOCUS CQ729566 1163 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 15500 from Patent WO02068579.
ACCESSION CQ729566
VERSION CQ729566.1 GI:42300960
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 15500 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
source 1..1163
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 35.8%; Score 743.8; DB 2; Length 1163;
Best Local Similarity 91.7%; Pred. No. 2.5e-186;
Matches 824; Conservative 0; Mismatches 12; Indels 63; Gaps 1;
Qy 300 GAACCTCTGCAATGTCCCATCAACCTCTCAGCTGCTCACTGAAAGAGGAGCAGGCCCA 359
Db 1 GAACCTCTGCAATGTCCCATCAACCTCTCAGCTGCTCACTGAAAGAGGAGCAGGCCCA 60

Qy 360 GTGAAGCACAGGAATGGACCCCCACCTGGCCCCACCCAAACCTGGACACGTTTACCC 419
Db 61 GTGAAGCACAGGAATGGACCCCCACCTGGCCCCACCCAAACCTGGACACGTTTACCC 120
Qy 420 CGGAGGAGCTGTGTCAGCAGATGAAGAGCTCTCTGACCAAGAAACCAACAGCTGAAAGAG 479
Db 121 CGGAGGAGCTGTGTCAGCAGATGAAGAGCTCTCTGACCGAAGAACCAACAGCTGAAAGAG 180
Qy 480 CCATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTTGAGAGCTTTCCGGCTGACAG 539
Db 181 CCATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTTGAGAGCTTTCCGGCTGACAG 240
Qy 540 AGAAACAGAGGAAGAACGCCAGTTTGTGATACAGAGCAAGCAAGCAAGAGCGTC 599
Db 241 AGAAACAGAGGAAGAACGCCAGTTTGTGATACAGAGCAAGCAAGCAAGAGCGTC 300
Qy 600 TAATGGCTTGAATCACTGAGAAATGAAGGAGAGCTTGGAAAACTTAAAGGGA 659
Db 301 TAATGGCTTGAATCACTGAGAAATGAAGGAGAGCTTGGAAAACTTAAAGGGA 360
Qy 660 AATCAGAAAGGTCTATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAGCGG 719
Db 361 AATCAGAAAGGTCTATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAGCGG 420
Qy 720 AGCAGAAAGAGCAGCTCAGGACCCAGGTGGTGAAGCTACAAGCAGAGAGGCGACCC 779
Db 421 AGCAGAAAGAGCAGCTCAGGACCCAGGTGGTGAAGCTACAAGCAGAGAGGCGACCC 480
Qy 780 TGTTCGGCATCTGTCTGAACTGACGCTCAAGCTCAAGCTGCTCCTCAGAGATTT 839
Db 481 TGTTCGGCATCTGTCTGAACTGACGCTCAAGCTGCTCCTCAGAGATTT 540
Qy 840 CTTTGTGTAATTAAGATGGCTGAAGGAGAGCAAGGGTCACTGTAAGAAATCAAGC 899
Db 541 CTTTGTGTAATTAAGATGGCTG----- 564
Qy 900 ATAGTCCTGGGGCCAGAGACAGTCTCCACTGGCAGGACATTTGTCTAATATAGGAGCA 959
Db 565 -----TGAGTTTTGGGCATTTGCTAAATATAGGAGCA 597
Qy 960 GATCTGCAGATGGGGCCAGAAATTAATTCGAACATGAGGAGTTAACTGTGAGCCAGCTCC 1019
Db 598 GATCTGCAGATGGGGCCAGAAATTAATTCGAACATGAGGAGTTAACTGTGAGCCAGCTCC 657
Qy 1020 TGTCTGTCTAAGGAAAGGAAATCAGAGGTGAGAGACTTGAAGTTGCACTCAAGAGG 1079
Db 658 TGTCTGTCTAAGGAAAGGAAATCAGAGGTGAGAGACTTGAAGTTGCACTCAAGAGG 717
Qy 1080 CCAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAGTAATCGTTCTGAGATTGAAACCC 1139
Db 718 CCAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAGTAATCGTTCTGAGATTGAAACCC 777
Qy 1140 ACACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTGTGA 1198
Db 778 ACACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTGTGA 836

RESULT 24
BD216639
LOCUS BD216639 710 bp DNA linear PAT 17-JUL-2003
DEFINITION Novel human genes and gene expression products ii.
ACCESSION BD216639
VERSION BD216639.1 GI:33026409
KEYWORDS JP 200251900-A/4781.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 710)
REFERENCE
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.S., Reinhard,C., Giese,K., Randazzo,F., Kennedy,G.C., Pot,D., Kassam,A., Lamson,G., Drmanac,R., Crkvenjakov,R., Dickson,M.,

Dmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V., Jones,L.W. and Crain,B.S.
Novel human genes and gene expression products ii
Patent: JP 2002519000-A 4781 02-JUL-2002;
CHIRON CORP.HYSEO INC
OS Homo sapiens (human)
PN JP 2002519000-A/4781
PD 02-JUL-2002
PF 28-JAN-1999 JP 2000556580
PR 28-JAN-1998 US 60/072910,24-FEB-1998 US 60/075954 PR
31-MAR-1998 US 60/080114,03-APR-1998 US 60/080515 PR
03-APR-1998 US 60/080666,21-OCT-1998 US 60/105234 PR
28-OCT-1998 US 60/105877
PI LOUIS T WILLIAMS,JAIME ESCOBEDO,MICHAEL A INNIS,PABLO PI
DOMINGUEZ GARCIA,
PI JULIE SUDDUTH KLINGER,CHRISTOPH REINHARD,KLAUSE GIESE,FILIPPO
PI RANDAZZO,
PI GIULIA C KENNEDY,DAVID POT,ALTAF KASSAM,GEORGE LAMSON,RADOJE
PI DRMANAC,
PI RADOMIR CRKVENJAKOV,MARK DICKSON,SNEZANA DRMANAC,IVAN LABAT,
PI DENA LESHKOWITZ,DAVID KITA,VERONICA GARCIA,LEE WILLIAM JONES,
PI BIRJIT STACHE CRAIN
PC C12N15/09,C12N15/09,C07K14/47,C07K14/82,C07K16/18,C12N1/15, PC
C12N1/19,
PC C12N1/21,C12N5/10,C12Q1/68,C12N15/00,C12N5/00,C12N15/00 CC
Novel human genes and gene expression products ii FH Key
Location/Qualifiers
FT source 1. .710 /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1. .710
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 32.0%; Score 665.2; DB 2; Length 710;
Best Local Similarity 98.7%; Pred. No. 2.3e-165;
Matches 681; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1009 GAGCCAGCTCCTGCTGTGCTTAAGGGAAGGGAATCAGAAGGTGGAGAGACTTCAAGTTGC 1068
DB 22 GATCCCTCGATTTCGTGTGCTTAAGGGAAGGGAATCAGAAGGTGGAGAGACTTGAAGTTGC 81
QY 1069 ACTCAAGGAGGCCCAAGAAAGAGTTTCAGATTTTGAAAAGAAAAACAAGTAATCGTTCTGA 1128
DB 82 ACTCAAGGAGGCCCAAGAAAGAGTTTCAGATTTTGAAAAGAAAAACAAGTAATCGTTCTGA 141
QY 1129 GATTGAAACCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGA 1188
DB 142 GATTGAAACCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGA 201
QY 1189 GACTGTTGGAAGGGAAGTGAAGCAGTCAACCTCCAGGTGACATCTCTGTTTAAGAGCT 1248
DB 202 GACTGTTGGAAGGGAAGTGAAGCAGTCAACCTCCAGGTGACATCTCTGTTTAAGAGCT 261
QY 1249 TCAAGAGGCTCATACAAAATTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAA 1308
DB 262 TCAAGAGGCTCATACAAAATTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAA 321
QY 1309 GTGTGAGGCCCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAAAGCAAGA 1368
DB 322 GTGTGAGGCCCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAAAGCAAGA 381
QY 1369 GCTTGTATTACTAAACAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAA 1428
DB 382 GCTTGTATTACTAAACAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAA 441
QY 1429 AATGGAACAGGCTTAAACAGAGGATGAAAAGTCCAAATTAACCTGCTCAGAGATGACACA 1488
DB 442 AATGGAACAGGCTTAAACAGAGGATGAAAAGTCCAAATTAACCTGCTCAGAGATGACACA 501
QY 1489 CAACAAGCTTCTTCAAGAACATAATAATGCAATTGAAAAACAATTTGAGGAACCTAACAGAAA 1548

Db 502 CAACAAGCTTCTTCAAGAACATAATAATGATGAAATGAAATGAGGAACTTAACAAGAAA 561
QY 1549 AGAGTCAGAAAAAGTGGACAGGCGCAGTGTGAGGAGACTGAGTCGAAAACTGGAACTGGC 1608
DB 562 AGAGTCAGAAAAAGTGGACAGGCGCAGTGTGAGGAGACTGAGTCGAAAACTGGAACTGGC 621
QY 1609 AGAAGAGGCTCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAAAACCCATTGCCAA 1668
DB 622 AGAAGAGGCTCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAAAACCCATTGCCAA 681
QY 1669 GCAGGAAGAGGAGCTGGAACCATGACCAT 1698
DB 682 GCAGGAAGAGG-CCTGGAACCATGACCAT 710
RESULT 25
AR770229
LOCUS AR770229
DEFINITION Sequence 4781 from patent US 6964868.
ACCESSION AR770229
VERSION AR770229.1 GI:83345508
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 710)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D.,
Sudduth-Klinger,J., Reinhard,C., Giese,K., Randazzo,F.,
Kennedy,G.C., Pot,D., Kassam,A., Lamson,G., Drmanac,R.,
Crkvenjakov,R., Dickson,M., Drmanac,S., Labat,I., Leshkowitz,D.,
Kita,D., Garcia,V., Jones,L.W. and Stache-Crain,B.
TITLE Human genes and gene expression products II
JOURNAL Patent: US 6964868-A 4781 15-NOV-2005;
Nuvelo, Inc.; Sunnyvale, CA
FEATURES
source Location/Qualifiers
1. .710
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 32.0%; Score 665.2; DB 2; Length 710;
Best Local Similarity 98.7%; Pred. No. 2.3e-165;
Matches 681; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1009 GAGCCAGCTCCTGCTGTGCTTAAGGGAAGGGAATCAGAAGGTGGAGAGACTTCAAGTTGC 1068
DB 22 GATCCCTCGATTTCGTGTGCTTAAGGGAAGGGAATCAGAAGGTGGAGAGACTTGAAGTTGC 81
QY 1069 ACTCAAGGAGGCCCAAGAAAGAGTTTCAGATTTTGAAAAGAAAAACAAGTAATCGTTCTGA 1128
DB 82 ACTCAAGGAGGCCCAAGAAAGAGTTTCAGATTTTGAAAAGAAAAACAAGTAATCGTTCTGA 141
QY 1129 GATTGAAACCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGA 1188
DB 142 GATTGAAACCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGA 201
QY 1189 GACTGTTGGAAGGGAAGTGAAGCAGTCAACCTCCAGGTGACATCTCTGTTTAAGAGCT 1248
DB 202 GACTGTTGGAAGGGAAGTGAAGCAGTCAACCTCCAGGTGACATCTCTGTTTAAGAGCT 261
QY 1249 TCAAGAGGCTCATACAAAATTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAA 1308
DB 262 TCAAGAGGCTCATACAAAATTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAA 321
QY 1309 GTGTGAGGCCCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAAAGCAAGA 1368
DB 322 GTGTGAGGCCCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAAAGCAAGA 381
QY 1369 GCTTGTATTACTAAACAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAA 1428
DB 382 GCTTGTATTACTAAACAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAA 441

QY	1429	AATGGAACAGGCTAAACACAGAGGATGAAAAGTCCAAATTAACTGTGCTACAGATGACACA	1488
Db	442	AATGGAACAGGCTAAACACAGAGGATGAAAAGTCCAAATTAACTGTGCTACAGATGACACA	501
QY	1489	CAACAAGCTTCTTCAAGACATAATAATCATTTGAAAACAATTTGAGGAACATAACAAGAAA	1548
Db	502	CAACAAGCTTCTTCAAGACATAATAATCATTTGAAAACAATTTGAGGAACATAACAAGAAA	561
QY	1549	AGAGTCAGAAAAAGTGGACAGGCGAGTGTGAAGGAACCTGAGTGAAGAAAACCTGGAACCTGGC	1608
Db	562	AGAGTCAGAAAAAGTGGACAGGCGAGTGTGAAGGAACCTGAGTGAAGAAAACCTGGAACCTGGC	621
QY	1609	AGAGAAGGCTGTGGCTTCCAAAACAGCTGCAAAATGGAATGAAGCAAAACCAATTGCCAA	1668
Db	622	AGAGAAGGCTGTGGCTTCCAAAACAGCTGCAAAATGGAATGAAGCAAAACCAATTGCCAA	681
QY	1669	GCAGGAGAGGACCTGGAACCATGACCAT	1698
Db	682	GCAGGAGAGGACCTGGAACCATGACCAT	710
RESULT 26			
AB050777		3853 bp	mRNA
LOCUS			linear
DEFINITION			Rattus norvegicus mRNA for injury-inducible-protein I-55, complete cds.
ACCESSION	AB050777		
VERSION	AB050777.1	GI:24475404	
KEYWORDS			Rattus norvegicus (Norway rat)
SOURCE			Rattus norvegicus
ORGANISM			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE			1. Myokai, F. and Oyama, M.
AUTHORS			Injury Inducible Gene I-55
TITLE			Published Only in Database (2002)
JOURNAL			2 (bases 1 to 3853)
REFERENCE			Myokai, F. and Oyama, M.
AUTHORS			Direct Submission
TITLE			Submitted (03-NOV-2000) Fumio Myokai, Nagasaki University Graduate School of Biomedical Sciences, Division of Oral Health Services Research; 1-7-1 Sakamoto, Nagasaki, Nagasaki 852-8588, Japan (E-mail:masataka@net.nagasaki-u.ac.jp, Tel:81-95-849-7663, Fax:81-95-849-7665)
FEATURES			Location/Qualifiers
source			1..3853
			/organism="Rattus norvegicus"
			/mol_type="mRNA"
			/db_xref="taxon:10116"
			/clone="i-55"
			/clone_lib="Rat liver lambda cDNA library (Stratagene)"
			202..1737
			/codon_start=1
CDS			/product="injury-inducible-protein I-55"
			/protein_id="BAC22658.1"
			/db_xref="GI:24475405"
			/translation="MWLFESPICRLCSGRKNTIYMPVCLYLCPRAWTOGCSHTRQGV CRSSPSALALIIGQPTKLPRLTSTNSCQIIQAAPFELGDSLADSLRLRFTQCTIL PSQPTFNHLSLLHLYDIGVLYEFLSHRRSRQLIQYKETWRPHLSFGSDAVHPS ASKAGPRYSKRKVPKHSYVVGCCFFSRSIDPEKANGHSALIEOTEGSTQKEE KDPESVGIEVETANQVASFELKLEAHLSEALMKRLQEKQALERNKNSATPS ELNEKELVSNKRLQVESMRSEIKMEQAKTEBKSLRLATLQATHLKLQHNKAL RTIEETKQQAQKVKVQLOELSEKLEAQLAKOLOMDSEMKOTIAKQEDLEMTA VRAQMEVTSDFHARAAREKHSEKEQLALQLAILKENNDPFDGSGRSQIMENQC RHGARTSDSDQAYLFORGAEDMSWHQGPQPSIPIHSCPKGCEVLDPDITLIQHVMD CII"
ORIGIN			
Query Match	31.1%;	Score 645.2;	DB 6; Length 3853;
Best Local Similarity	80.7%;	Pred. No. 3.2e-160;	
Matches 809; Conservative	0;	Mismatches 173;	Indels 20; Gaps 4;

RESULT 27
AF389351
LOCUS

AF389351 2476 bp mRNA linear VRT 16-AUG-2002

DEFINITION Gallus gallus FIP-2 (FIP-2) mRNA, complete cds.
ACCESSION AF389351
VERSION AF389351.1 GI:21666286
KEYWORDS Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 2476)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
TITLE FIP-2, an IkappaB-kinase-gamma-Related Protein, Is Associated with the Golgi Apparatus and Translocates to the Marginal Band during Chicken Erythroblast Differentiation
JOURNAL Exp. Cell Res. 278 (2), 133-145 (2002)
PUBLISHED 12169269
REFERENCE 2 (bases 1 to 2476)
AUTHORS Stroissnigg, H., Repitz, M., Miloloza, A., Linhartova, I., Beug, H., Wiche, G. and Propst, F.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2001) Institute of Biochemistry and Molecular Cell Biology, University of Vienna, Dr. Bohrgasse 9, Vienna A-1030, Austria

FEATURES

source

1. .2476
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /db_xref="taxon:9031"
 /cell_line="HD-3"
 1. .2476
 /gene="FIP-2"
 65. .1735
 /gene="FIP-2"
 /note="IKKgamma-related protein; unknown function"
 /codon_start=1
 /product="FIP-2"
 /protein_id="AAM73640.1"
 /db_xref="GI:21666287"
 /translations="MSSKQIPRAENGHCRRSMGMDSPPTLTSTYTPRMVQOM
 KELITNNELKEAMKLHNOAKDRVELSIWEKQKEEFVETKEKAQCLLAKV
 ENFQLOQQLSKEREGAMECGATPEKARQLKSKVRLQREKADLLAIISELQVK
 LVNALSDFEVLGMNEVNRITARENDSSMASNIAVIIRSKSADESKNLESLVT
 SOLLCLRNERTQREKLEQDHKRLSKMENETSNCLSGTQTNQEESEATGS
 EYSELKKQICALFKELQEAHEKLEAELIQKLEQKQCTLEKVNAAAELSEKQOLI
 YTIKLEQVSEVQAEVLEQAKTODEKTRYSSLODAYNKLAELEAMKTI SEMKVK
 EHRDVKVVEELNAKVLLEQALAKQLQMDENKQLIAKQBEDLETMAVLRAQHEVY
 CSDFHARAAREKIHEKEQLAVQLAYLLKQONLEDLGRSSLAENQNRHGRAPDRE
 HSPRLVQRTGSGQEWPEQRNISIYSCPKCEBILPDLDLTQIHVMDCIN"

ORIGIN

Query Match 31.0%; Score 644.4; DB 11; Length 2476;
 Best Local Similarity 63.2%; Pred. No. 5.8e-160;
 Matches 1106; Conservative 0; Mismatches 581; Indels 63; Gaps 5;
 QY 297 CAGGAATCTGTGCAATGTCCCACTCAACCTCTCAGCTGCCTCACTGAAAAGGAGGACAGCC 356
 DB 51 CGGTCACTGCAGAGATGTCAGCAAGCCGAGATCCGTCTGTGAAATGGAGACATT 110
 QY 357 CCAGTGAAGACAGAGAAATGACCCCCCACTGCGCCCAACCCAAACCTTGGACACGTTTA 416
 DB 111 GCAGGAGTAAATGGAGAACGGAAATGGACAGCATGCTCCGCCCACTCAGTACCTACA 170
 QY 417 CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGACCACTGCAAG 476
 DB 171 CACCCGAAGAGATGGTACAGCAGATGAAGAGCTGATCACTGAGAACTATGATGTTAAAG 230
 QY 477 AAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTTGAGGAGCTTTCGGCTGGA 536
 DB 231 AAGCGATGAATGGACACACCAAGCAATGAAGACCGCTATGAGGAGCTTTCATCTGGA 290
 QY 537 CAGAGAAA CAGAGGAAGAACCGCAGTCTTTTTCAGATACAGAGCAAGAAAGCAAGAGC 596

Db 291 GAGAGAGCAAAAAGGAGAGCGAGAGTTCACGAGACAAAAGTTTAAAGAAAGCCCAAGCAGT 350
 QY 597 GTCTAATGGCCTTTGAGTCATGAGAAATTTGAGGAGAGCTTTGGAATAACTTAAAG 656
 Db 351 GCTTGTAGCAAGTGGTGTGAAATAGCAGCTGCAGCAACCACTTTCAGAGCTTAAAGG 410
 QY 657 GGAATACTAGAAAGGTCAATCTTGAGAGCCCACTGATGACTCCAGGCTTCCAGGCGCCGAAG 716
 Db 411 AAAGAGAGAAAGGAGCTGAAATGGA-----GGGCTGTGCAACTC 449
 QY 717 CGGAGCAGGAAAAGGACCAGCTCAGACCCAGGTGTGAGGCTTCAAGCAGAGAGAGGAG 776
 Db 450 CTGAGAAGGAAAGCAGGCGAGCTCAAGTCCAAGTGCAGCGGCTCCAGGCTCAGAAGGCGAG 509
 QY 777 ACCTGTGGGCATCTGTGTCTGAAGTGCAGCTCAAGCTGAACTCCAGCGGCTCCTCACAAG 836
 Db 510 ATCTGTGGCCATCATTTTCAGAGCTGCAGGTCAAGCTGAA---CATTGCCTCAGCAGAG 566
 QY 837 ATTCTTTGTGTAATTAGGATCGCTGAAGGAGAGCAGAGGCTCAGTAAAAAGAAATCA 896
 Db 567 ACTCTTTGTGGAGATTGGGATGAACGAGAAAGTAAATCGAATCGAAGAGAGAAATCAAG 626
 QY 897 AGCATAGTCTCTGGGCCACGAGAACAGTCTTCACTGGCAGCGGATTTGTCTAAATATAGGA 956
 Db 627 ATAATAGCAGTGAATGGCCAGTAACATCGCTGTCTACA-----TCAGGA 671
 QY 957 GCAGATCTGCAGATGGGCCCAAGAAATTTACTTCGAAATGAGGAGTTAACTGTGAGCCAGC 1016
 Db 672 GCAATCTGCAGATGAGTCCAAGAT---TTGGAATCTGAGGAGCTGACTGTGAGCCAGT 728
 QY 1017 TCCTGCTGTCCCTAAGGGAAGGGAATCAGAAGGTGGAGAGACTTGAAGTTTGCCTCAAGG 1076
 Db 729 TGCTCTGCTGCCTTAGAATGAACTCAGAGGAGGAGAAAATCTTGAGAGGAGCTGCAGG 788
 QY 1077 AGCCAAAGAAAGAGTTTCAGATTTTGAAGAAAAACAAGTAAATCGTTCTGAGATTGAAA 1136
 Db 789 ATCAAAAGAGAGACTTTCAAAGATGGAGAAATGAAACAAGCAACTGCTCGAAAGCGGA 848
 QY 1137 CCCAGACAGAGGGGACACAGAGAAAGAGATGATCAAGAGAAAGGCCCGGAGACTGTTG 1196
 Db 849 CGCAAAAC-----TAAACGAGAAAGAGAGAGTTTTCAGAGGCTATTG 887
 QY 1197 GAAGCGAAGTGAAGACACTGAACCTCCAGGTGCATCTCTGTTTAAAGGAGCTTCAAGAGG 1256
 Db 888 GCAGTGAAGTAGAAGTCTGAGAGCAAAATTTGTGCCCTCTTTAAAGAGCTCCAAGAGG 947
 QY 1257 CTCATACAAAATCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTGAG 1316
 Db 948 CCCACGAGAAGTTAAAGAAAGCAGAACTGATTTCAGAAAGAACTTCAAGAAAGAGTGC 1007
 QY 1317 CCCTTGAAGGAAAATTTCTGCAATTCATCAGAGTTGAATGAAAGCAGAGAGCTTGTGTT 1376
 Db 1008 CATTTGAAAAAGTAAATTCAGCTGCTGCAACCGAAATTGAGGAGAGAGCAGCTCATAT 1067
 QY 1377 ATACTAAACAAAAGTTAGAGCTTACAAGTGAAGAGATGCTATTCAAGAAATCAAAATGAA 1436
 Db 1068 ACACCATCAAGAGCTGGAGCTTCAGTGGAGAGGCTGCAGGAGAGCTCAGAGCTAGAAC 1127
 QY 1437 AGGCTAAACACAGAGATGAAAAGTCCAAATTAATCTGTGTCTCAGATGACACACAAAGC 1496
 Db 1128 AAGCCAAAGACACAGGACGAAAAGACAAGATACAGTAGCTTGCAGGATGCTTTCAACAAAC 1187
 QY 1497 TTCTTTCAAGACATTAATGCAATTTGAAAAACAATTTGAGGAACTAACAGAAAAAGAGTCA 1556
 Db 1188 TCCTTTGCTGAATCTCACTGAAGCAATGAAAACAATCAGTGAAATGAAAGTCAAGAGCATG 1247
 QY 1557 AAAAAGTGGACAGGCGAGTGTGGAAGGAACCTCAGTGAAGAACTGGAACCTGGCAGAGAG 1616
 Db 1248 ACAGGCTGATGAAGTTGTGTGGAGGAACCTGATGCAAGAGGTGTTGTGGCAGAGACAAG 1307
 QY 1617 CTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCAAGAGGAG 1676
 Db 1308 CACTTGTCTGCAAGAGCAGCTCCAGATGGATGAATGAGCAGCTAATTTGCTTAGCAGAGG 1367

QY	1677	AGGACCTGGAACCAATGACCATCTCCAGGGCTCAGATGAAAGTTTACTGTTCTGATTTTC	1736
Db	1368	AGGATCTTTGAAACCATGGCTGCTCGTCTCAGATGGAGGTTTATTGTTCTGATTTCC	1427
QY	1737	ATGCTGNAAGACGCGAGAGAAAATTCATGAGGAAGAGCAACTGGCAATTCGACG	1796
Db	1428	ATGCCGAGAGGCGACGAGAGAGATCCATGAGGAGAAGAGAGCTGGCTGTCACG	1487
QY	1797	TGGCAGTTCTGCTGAAAGAGAATGATGCTTTTGAAGACGGAGCGAGGCACTCTTGATGG	1856
Db	1488	TGGCATACCTGCTAAAGAGAGCAAAACCTTGAAGATCTTGGCGAAGCTCTTTGGCAG	1547
QY	1857	AGATGACAGTCTCATGCGGGCGAGAACTGATGATCTTGACAGAGGCTTACCTTGTC	1916
Db	1548	AGATGAGAATCGCCACGAGCGAGGACACAGATCGGAGCACTCACTCGCTTGTCC	1607
QY	1917	AAAGAGGAGCTGAGNACAGGAGTGGCGGCACAGCGGAGATATTCGATTCATCTGTC	1976
Db	1608	AAAGAGGAATCGGTAGTCAAGAATGGCCAGAACAGCGGAATATCAATATATCTCTGC	1667
QY	1977	CAAAGTGTGAGAGGTTCTGCTGACATAGACAGTTACAGATTACAGTGTGATGGA	2036
Db	1668	CAAAATGGAAGAAATCTGCCAGACCTGCAGACCGCTGCAGATCCATGTTATGGACTGA	1727
QY	2037	TCATTTAAGT 2046	
Db	1728	TTAATTGAAT 1737	
RESULT 28	AF380358	2555 bp mRNA linear	VRT 12-FEB-2002
LOCUS	Gallus gallus AG9C5 mRNA, complete cds.		
DEFINITION	AF380358		
ACCESSION	AF380358.1	GI:14579552	
VERSION			
KEYWORDS	Gallus gallus (chicken)		
ORGANISM	Gallus gallus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;		
TITLE	Phasianinae; Gallus.		
JOURNAL	1 (bases 1 to 2555)		
PUBMED	Li, B. and Gallin, W. J.		
REFERENCE	Differential localization of chicken FIP2 homologue, Ag-9C5, in		
AUTHORS	secretory epithelial cells		
TITLE	Exp. Cell Res. 272 (2), 135-145 (2002)		
JOURNAL	11777338		
PUBMED	2 (bases 1 to 2555)		
REFERENCE	Gallin, W. J.		
AUTHORS	Direct Submission		
TITLE	Submitted (10-MAY-2001) Department of Biological Sciences,		
JOURNAL	University of Alberta, Z-631, Edmonton, Alberta T6G 2E9, Canada		
FEATURES	Location/Qualifiers		
source	1..2555		
	/organism="Gallus gallus"		
	/mol_type="mRNA"		
	/db_xref="taxon:9031"		
5'UTR	1..37		
CDS	38..1708		
	/codon_start=1		
	/product="Ag9C5"		
	/protein_id="AAK69306.1"		
	/db_xref="GI:14579553"		
	/translation="MSSKQIRPAENGHCRSMKMGDSMAPPTLSTVTPBEMVOOM		
	KELITENNELKAMKLHNQAKMDRYEELSIWREKQKEEFVETKFEAKQCLLAKCV		
	ENQLQQQLSKLEREGEAMEGCTPEKEARQLKQVRLQRLQKRAADLALISELOVK		
	LNIASAEDEFVETGMNEVNRARENQDSSNSNAIVIRSKSADESKNLESELT		
	VSOLLCLRNERTORREKLEKLDQHKERLSKMNENETSNLCSTQTNQBESESAIGS		
	EVELSKQICALFVKLEQAEHLKAEELIQKLEKQCTLEKVNLSAAATELBKQOLI		
	YTIKLEIQVES/QVAEKLQEAOKTQDEKTRYSSLODAYNKLKLAELTEAMKTTISEMKVK		
	EHDVDKVVVEELNVLAEQALAAKQLOMDKMLIAKQEDLETMLVLAQMEVY		
	CSDFHARAAREKIHEKEQLAVQLAYLLKEQNLEDIGRSSLAENQRHGAPRDE		

3'UTR	ORIGIN	Query Match	30.7%; Score 637.6; DB 11; Length 2555;
		Best Local Similarity	63.0%; Pred. No. 3.7e-158;
		Matches 1101; Conservative	0; Mismatches 584; Indels 63; Gaps 5;
QY	297	CAGGAATCTTGCATATGTCCTCAATCAACTCTCAGTGTCTCATCTGAAAGAGGACAGCC	356
Db	24	CGGTCACTGCAGAGATGTCAGCAAGCCGAGATCCGTCTCTGCTGAAATGAGACATTT	83
QY	357	CCAGTGAAGCACAGGAATGGACCCCGCCACCTGGCCACCAACCTGGACACGTTTA	416
Db	84	GCAGGAGTAAATGGAGAACGGAATGGACAGCATGGCTCCCCCACCCTCAGTACCTACA	143
QY	417	CCCCGGAGGAGCTGCTGCAGCAGATGAAAAGAGCTCTGTACCAAGAAACCAACAGCTGAAAG	476
Db	144	CACCCGAGAGATGTTGACAGCAGATGAAAAGAGCTCATCTAGAACATGAGTTAAAG	203
QY	477	AAGCCATGAAGCTAAATTAATCAAGCCATGAAAGGAGATTTGAGAGCTTTCCGGCTGGA	536
Db	204	AAGCATGAAGTTGCAAAACCAAGCAATGAAGACCGCTATGAGGAGCTTTCCATCTGGA	263
QY	537	CAGAGAACAGNAGGAAGAACCCCAAGTTTGTGATACAGACAGCAAGAACGAAAGAGCC	596
Db	264	GAGAGAACAAAGGAAGAGCGAGATTTTACGACAAAGTTTAAAGAACGCAAGCAGT	323
QY	597	GTCTAATGGCTTGAATCATGAGAATGAGAAATTTGAAGGAAGAGCTTTGAAAACTAAAAAG	656
Db	324	GCTTCTAGCCAGTGGTGGTTGAAATGAGCAGCTGCAGCAACAATCTCAGAGCTTAAGG	383
QY	657	GGAAATCAGAAAGGTTCATCTGAGGACCCCACTGATGATCTCCAGGCTTCCAGGGCCGAAG	716
Db	384	AAAGAGAAGAGGAGCTGAATGGA-----GGGCTGTGCAACTC	422
QY	717	CGGACAGGAAAGAACCCAGCTCAGGACCCAGGTGGTGAAGGCTACNAGCAGAGAGGCGAG	776
Db	423	CTGAGAAAGAAAGCGAGGAGCTCAAGTCCAGGTCCAGGCTCCAAAGCTGAGAAAGGCGAG	482
QY	777	ACCTGTGGGACCTGCTCTGAATCTGACCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAG	836
Db	483	ATCTGTGGCCATCATTTTCAGAGCTGCGAGGTCAAGCTGAA---CATTTGCTCAGCAGAG	539
QY	837	ATTCCTTTGTTCAAAATTAGGATGGCTGGAAGGAAGAGCAAGAGGTTCAGTAAAGAAATCA	896
Db	540	ACTCCTTTGTGGAGATTGGATGAAACGAAGAAAGTAAATCGAACTGCAAGAGAAATCAAG	599
QY	897	AGCATAGTCTGGGCCCCACGAGAACAGTCTCCACTGGCAGGCAATGTCTAAATATAGGA	956
Db	600	ATAATAGCAGTGAATGGCCAGTAAACATCGCTGTCTACA-----TCAGGA	644
QY	957	CGAGATCTGCAGATGGGGCCAAAGAAATTAATCTCGAATCAGGAGTTAACTGTGAGCCAGC	1016
Db	645	GCAATCTGCAGATGAGTCCAAGAT--TTGGAATCTGAGGAGCTGACTGTGAGCCAGT	701
QY	1017	TCCTGCTGTCCCTAAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG	1076
Db	702	TGCTCTGCTGCTTAGAAATGAAATCTCAGAGGAGGGAAGAACTTGAGAAAGGAGCTGCAGG	761
QY	1077	AGGCCAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAAAGTAACTCTCTGAGATTGAAA	1136
Db	762	ATCACAAAGAGAGACTTTTCAAGATGGAGAAATGAAACAAGCAACTCTCCCTGGAAAGCGGA	821
QY	1137	CCACAGACAGAGGGGAGCACAGAGAAAGAGATGATGAAGAGAAAGGCCCCGGAGACTGTTG	1196
Db	822	CGCAAC-----TAAACGAGAAAGAGAGAGTTTCAAGAGGCTATTG	860
QY	1197	GAAGCGAAGTGAAGACATGGAACCTTCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGG	1256
Db	861	GCAGTGAAGTAGAAGTCTGAAGAGCAAAATTTGTGCCCTCTTTTAAAGAGCTCCAAGAGG	920
QY	1257	CTCATACAAACTCAGCGAAGCTGAGCTAATGAAGAAAGAGACTTTCAAGAAAAAGTGTGAGG	1316

```
Db 921 CCCACGAGAGTTTAAAGAACGACAGAACTGATTTCAGAGAAAGAACTTCAAGAAAGAGTGCCAAA 980
Qy 1317 CCCTTTGAAAGGAAAAATTTCTCAATTCATCAGAGTTGAATGAAAAAGCAAGAGCTTGTTT 1376
Db 981 CATTTGGAAGAAAGTAAATTCAGTCTGCAACGGAAATTTGGAGGAGAGCAGCAGCTCATAT 1040
Qy 1377 ATACTAACAAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC 1436
Db 1041 ACACCATCAAGAACCTGAGCTTCAGGTGGAGAGCGTGCAGGCAGAGGTCAAGCTAGAAC 1100
Qy 1437 AGGCTAAAAACAGAGGATGAAAGTCCAAATTAATTAACGTGCTACAGATGACACACAAAGC 1496
Db 1101 AAGCCACAGACAGGAGAAAGACAGATACAGTAGCTTGCAGGATGCTTACAAACAAC 1160
Qy 1497 TTCTTTCAAGACATATAATGCAATGAAAAAATTTGAGGAACTTAAACAGAAAAAGAGTCAG 1556
Db 1161 TCCTTGCTGAACTCACTGAGCAATGAACAAATCAGTGAATGAAGTCAAGAGGATG 1220
Qy 1557 AAAAAGTGGACAGGCGAGTGTGAAGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGG 1616
Db 1221 ACAGGTGGATAAAGTTGTGTGGAGAACTGAATGCAAAAGTGTGTGTGGCAGAACAG 1280
Qy 1617 CTCTGGCTTCCAAACAGCTCAATGATGAATGAATGAACAACTGGAACCTGGCAGAGGAG 1676
Db 1281 CACTTGCTGCAAAAGCAGCTCCAGATGATGAATGAATGAACGAGCTAAATTTGCTAAGCAAGG 1340
Qy 1677 AGGACCTTGAAGAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTC 1736
Db 1341 AGGATCTTGAACCATGGCTGTCTCGTCTCAGATGGAGGTTTATTTCTTGACTTCC 1400
Qy 1737 ATGCTGAAAGCAGCAGAGAGAGAAATTCATGAGGAAAGAGCAACTGGCAATGCGAC 1796
Db 1401 ATGCCGAGAGGCGGCCAGAGAAAAATCCATGAGGAGAGGAGCAGCTGGCTGTCCAAC 1460
Qy 1797 TGGCAGTTCTGCTCAAAAGAGATGATGCTTCGAGAGGAGGAGGAGCAGCTCTTGATGG 1856
Db 1461 TGGCATACCTGCTTAAAGAGACGACAAACCTTGGGAAGATCTTGGCCGAAGCTCTTTGGCAG 1520
Qy 1857 AGATGACAGAGTCTGTCATGGGCGAGAACAAAGTGAATCTGACCTGACAGCAGGCTTACCTTGTC 1916
Db 1521 AGATGACAGACCGCCAGAGCCAGGACCAAGATCGGAGCACTACCTGCGCTTGTC 1580
Qy 1917 AAAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTTCCGATTCATTCTGTC 1976
Db 1581 AAAGAGAACTGGTAGTCAAGATGGCCAGAACAGCGGAATATCTCAATATATTTCTGTC 1640
Qy 1977 CCAAGTGTGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCAGCTGATGATGCA 2036
Db 1641 CAAAATGTGAAGAAATTTCTCCAGACCTGGACAGCTGCAGATCCATGTTATGAGCTGCA 1700
Qy 2037 TCATTTAA 2044
Db 1701 TTAATTGA 1708

RESULT 29
AB222073
LOCUS AB222073 1496 bp mRNA linear ROD 21-JAN-2006
DEFINITION Rattus norvegicus FIP-2A variant mRNA for 14.7K-interacting
protein-2 type A variant, partial cds.
ACCESSION AB222073
VERSION AB222073.1
KEYWORDS AB222073.1 GI:85658722
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1
AUTHORS Myokai, F.
TITLE FIP-2A variant
JOURNAL Unpublished
```

```
REFERENCE 2 (bases 1 to 1496)
AUTHORS Myokai, F. and Oyama, M.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2005) Fumio Myokai, Okayama University Graduate
School of Medicine, Dentistry, and Pharmaceutical Sciences,
Periodontal Science, Shikata-cho 2-5-1, Okayama 700-8525,
Japan (E-mail: myokai@md.okayama-u.ac.jp,
Tel:81-86-235-6677 (ex.6678), Fax:81-86-235-6679)
FEATURES
Location/Qualifiers
source
1..1496
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="liver"
/clone_lib="Rat liver Marathon-Ready cDNA kit"
/country="USA: CA, Palo Alto"
1..>1496
/gene="FIP-2A variant"
746..>1496
/gene="FIP-2A variant"
/codon_start=1
/product="14.7K-interacting protein-2 type A variant"
/protein_id="BAE78454.1"
/db_xref="GI:85658723"
/translation="MKRLQEKQALERNKNSATPSELNKEQLVYNRKLELOVESMR
SEIKMEQAKTEEEKSLRLATQATHDKLQHLNKLARTIELTKQAEKVDKVALEQELS
EKLELAQALQKQLOMDKQTIKQEBDELTMVLAQMEVYVCSDFHAEARAAREKI
HEKELOALQALAILKENNDFEDGSGRSQIMEMQCRHGARTSDSDQQAAYLFRGAEDM
SWHQGQPRSPRIPIHSCPKEVLPDITLQIH"
ORIGIN
Query Match 30.3%; Score 629.8; DB 6; Length 1496;
Best Local Similarity 81.3%; Pred. No. 5e-156;
Matches 772; Conservative 0; Mismatches 162; Indels 15; Gaps 3;
Qy 1088 AGAGTTTCAGATTTTGAAGAAAAACAAGTAATGCTTCTGAGATTTGAAACCCAGACAGAG 1147
Db 551 AGAATTTTCAGATTTTGAAGAAAAAGCAACCGCCATTTCTCGATTGAGACCCAGACAGAG 610
Qy 1148 GCGAGCACAGAGAAAGAGATGATGAGAGAAAGGCCGAGAGCTTCAAGAGGCTCATACAAA 1207
Db 611 GGGAGCACACAAAAG---AAGAGGAGGACAAAGACCCAGAGAGTGTGGAAATCGAAAGT 667
Qy 1208 GAAGCACTGAACCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAA 1267
Db 668 GAACCTTGACCTTCAAGTGGCTCTCTGTTTAAAGGAGCTTCAAGAGGCGCACACAAG 727
Qy 1268 CTGAGGAGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTGAGGCCCTTGAAAGG 1327
Db 728 CTGAGTGAGGCCGAGCTGATGAAGAGAGAGACTTCAAGAAAAAGTGTGAGGCTCTGGAAGG 787
Qy 1328 AAAAAATTCGATTTCCATCAGAGTTGAAATGAAGAGAGAGCTTGTATTATACAAA 1387
Db 788 AAGAACTCTGCAACCCCATCGGAGCTGAATGAAGAAAGCAAGAGCTCGTTTACAGTAACAG 847
Qy 1388 AAGTTAGAGCTACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACA 1447
Db 848 AAGTTAGAGCTGAGGTGGAGAGCATGCGCTCGAAATCAAGATGGAGAGGCCACAGACA 907
Qy 1448 GAGGATGAAAGTCCAAAATTAAGTGTGCTACAGATGACACACAAAGCTTCTTCAAGAA 1507
Db 908 GAGGAGGAGAAATCCAGGTTAGCCACTCTTCAGGCAACACACAGCAAGCTTCTTCAGGAA 967
Qy 1508 CATAATAATGCATTGAAAAAATTTAGGAACTTGAAGAACTTGAAGAAAGAGTCAAGAAAAAGTGGAC 1567
Db 968 CACAATAAGCTCTGAGAACAAATTTGAAGAACTTAACCAAAACAAAGCAGCAAAAAAGTGGAC 1027
Qy 1568 AGGGCAGTGTGTAAGAACTGAGTGAAGAACTCGAACTGCGAGAGAGGCTCTGGCTTCC 1627
Db 1028 AAGTGCAGCTGCAGAGGCTCAGCGAAGAGCTGGAGCTGGCGGAGCAGGCTCTGGCGTCC 1087
Qy 1628 AAACAGCTGCAATGGATGAATGAAGCAAAACATTTGCCAAGCAGGAAGAGGACCTGGAA 1687
```

Db	1088	AAGCAGCTCCAGATGATGAGATGAAGCAGACGATCGCCCAAGCAGGAGGAGGACCTGGAG	1147	PUBMED REFERENCE AUTHORS	14702039
Qy	1688	ACCATGACCATCTCAGGCGCTCAGATGGAAGTTTACTGTTCTGATTTTTCATGCTGGAAGA	1747	Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T., Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R., Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiya,H., Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.	
Db	1148	ACCATGGCGCTCTCAGGCGCTCAGATGGAGGTGTACTGTTTCAGATTTCCATGCTGAGAGG	1207	Unpublished NEDO human cDNA sequencing project	
Qy	1748	GCAGCGAGAGAGAAATTCATGAGGAAAGGAGCACTGGCATTTGGCAGTGGCAGTTCTG	1807	Isogai,T. and Yamamoto,J.	
Db	1208	GGCGCAAGGAGAGATCCATCAGGAAAAAGGACGCTGGCGCTTGCGACTCGCCATTTTG	1267	Direct Submission	
Qy	1808	CTGAAGAGAAATGATCTTTCGAACGCGAGC---AGCAGTCTCTTGATGATGAGATGCGAG	1864	Submitted (15-JUL-2003) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: flij-cdna@nifty.com, Tel:81-438-52-3975, Fax:81-438-52-3986)	
Db	1268	CTGAAGAGAAACAATGACTTTGAAGATGGAGCGCAGTAGCGAGTCTTGATGAAGAAATCGAG	1327	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: Reverse Proteomics Research Institute, HRI and RAB.	
Qy	1865	AGTCGTGTCGGCGGAGAAACAGTGAATCTGACCGACGCGCTTACCTGTTTCAAGAGGA	1924	Location/Qualifiers 1. 1913 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="NOVAR2002039" /tissue_type="ovary" /clone_lib="NOVAR2" /dev_stage="adult" /note="cloning vector: pME18SFL3"	
Db	1328	TGCCGCGCAGGGCGGAGAACCGAGCTCTGACCGACGCGCTTACCTGTTTCAAGAGGA	1387	FEATURES source	
Qy	1925	GTGAGGACAGGGACT-----GGCGGCAACAGCGGAATATTCCGATTCATTCCTGTC	1975	ORIGIN	
Db	1388	GCGCGAGGACATGAGCTGGCAGCATGGCGACGACGCCGCGAGTATTCCCATCTCACTGTCG	1447	Query Match 29.9%; Score 620.2; DB 5; Length 1913; Best Local Similarity 99.5%; Pred. No. 1.7e-153; Matches 622; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Qy	1976	CCCAAGTGGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCACG	2024	Qy 1453 TCAGAAAGTCCAAATTAACTGCTGTCTACAGATGACACACAAAGCTTCTTCAAGAACATAA	1512
Db	1448	CCCAAGTGGGAGGTCTGCGGACATTTGACGCGCTTCAGATACATG	1496	Db 1268 TAATAGGTCCAAATTAACTGCTGTCTACAGATGACACACAAAGCTTCTTCAAGAACATAA	1327
RESULT 30				Qy 1513 TAATGCAATTGAAAAACAATTTGAGGAACTTAAACAAGAAAGAGTCAAGAAAGTGACAGGCG	1572
AK122746				Db 1328 TAATGCAATTGAAAAACAATTTGAGGAACTTAAACAAGAAAGAGTCAAGAAAGTGACAGGCG	1387
LOCUS				Qy 1573 AGTGTGTAAGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCTGGCTTCCAAACA	1632
DEFINITION				Db 1388 AGTGTGTAAGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCTGGCTTCCAAACA	1447
ACCESSION				Qy 1633 GCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAGGACCTGGAACCAT	1692
VERSION				Db 1448 GCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAGGACCTGGAACCAT	1507
KEYWORDS				Qy 1693 GACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTTCATGCTGGAAGAGCAGC	1752
SOURCE				Db 1508 GACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTTCATGCTGGAAGAGCAGC	1567
ORGANISM				Qy 1753 GAGAGAGAAAAATTCATGAGGAAAAAGGAGCAACTGGCAATTCAGCTGGCAGTTCTGCTGAA	1812
Homo sapiens				Db 1568 GAGAGAGAAAAATTCATGAGGAAAAAGGAGCAACTGGCAATTCAGCTGGCAGTTCTGCTGAA	1627
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				Qy 1813 AGAGAAATGATGCTTTTCGAGAGCGGAGCGGAGTCTTGTGATGGAGATGCAGAGTCTGCTCA	1872
REFERENCE					
AUTHORS					
Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,M., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hirakawa,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hottuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T.O., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shionata,N., Sano,S., Moriya,S., Momiya,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Sato,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.					
Complete sequencing and characterization of 21,243 full-length human cDNAs					
TITLE					
JOURNAL					

```
Db 1628 AGAAGATGATGCTTTCCGAAGACGGAGCGAGCTCCTTGTATGAGATGCAGAGTCGTCA 1687
Qy 1873 TGGGGCGAGAACAGTCACTCTGACACGAGGCTTACCTTGTTCCAAAGAGGAGCTGAGGA 1932
Db 1688 TGGGGCGAGAACAGTCACTCTGACACGAGGCTTACCTTGTTCCAAAGAGGAGCTGAGGA 1747
Qy 1933 CAGGACTGGGGCAACAGCGGAATATTCGGAATTCATTCCTGCCCCCAAGTGTGGAGAGGT 1992
Db 1748 CAGGACTGGGGCAACAGCGGAATATTCGGAATTCATTCCTGCCCCCAAGTGTGGAGAGGT 1807
Qy 1993 TCTGCTGCATAGACACGATTACAGATTACAGTTACGTTGATGATTCATTTAAAGTGTGAT 2052
Db 1808 TCTGCTGCATAGACACGATTACAGATTACGTTGATGATTCATTTAAAGTGTGAT 1867
Qy 2053 GTATCACTCCCCAAAACCTGTGCT 2077
Db 1868 GTATCACTCCCCAAAACCTGTGCT 1892

RESULT 31
CS071068/c
LOCUS 3454 bp DNA linear PAT 05-MAY-2005
DEFINITION Sequence 231 from Patent WO2001032927.
ACCESSION CS071068
VERSION CS071068.1 GI:63088472
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
JOURNAL Hominiidae; Homo.
1
Sornasse,T., Seilhamer,J.J. and Watson,G.A.
Tissue specific genes of diagnostic import
Patent: WO 2001032927-A 231 10-MAY-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
location/Qualifiers
1..3454
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 25.4%; Score 526.6; DB 2; Length 3454;
Best Local Similarity 99.1%; Pred. No. 1.2e-128;
Matches 540; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 1307 AAGTGTGAGCCCTTGAAGGAAATTTCTGCAATTCATCAGATTGAATGAAAGCAA 1366
Db 2549 AGGTGTGAGCCCTTGAAGGAAATTTCTGCAATTCATCAGATTGAATGAAAGCAA 2490
Qy 1367 GAGCTGTTTATCTAACAAAGTTAGAGCTCAAGTGAAGAGCATGCTATCAGAAATC 1426
Db 2489 GAGCTGTTTATCTAAC-AAAAGTTAGAGCTCAAGTGAAGAGCATGCTATCAGAAATC 2431
Qy 1427 AAAATGGAACAGGCTAAACAGAGGATGAAAGTCCAAATTTAACTGTGCTACAGATGACA 1486
Db 2430 AAAATGGAACAGGCTAAACAGAGGATGAAAGTCCAAATTTAACTGTGCTACAGATGACA 2371
Qy 1487 CACAACAGGCTTTCAAGAACATAATAATGCAATTGAAACCAATGAGGAACTTAAACAAGA 1546
Db 2370 CACAACAGGCTTTCAAGAACATAATAATGCAATTGAAACCAATGAGGAACTTAAACAAGA 2311
Qy 1547 AAAGAGTCAGAAAGTGGACGGGAGTCTGTAAGAACTGATGAAAACTGGAACTG 1606
Db 2310 AAAGAGTCAGAAAGTGGACGGGAGTCTGTAAGAACTGATGAAAACTGGAACTG 2251
Qy 1607 GCAGAGAGGCTCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAAACTTGCC 1666
Db 2250 GCAGAGAGGCTCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAAACTTGCC 2191
Qy 1667 AACGGAAGAGACCTTGAAACCAATGACATCCTCAGGCTCAGATGGAAGTTTACTGT 1726
|||||
```

```
Db 2190 AAGCAGGAAGAGACCTGGAAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGT 2131
Qy 1727 TCTGATTTTTCATCTGTAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGGAGCAACTG 1786
Db 2130 TCTGATTTTTCATCTGTAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGGAGCAACTG 2071
Qy 1787 GCATTGAGCTGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTCGAAGACGGAGCGAGCAG 1846
Db 2070 GCATTGAGCTGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTCGAAGACGGAGCGAGCAG 2011
Qy 1847 TCCTTT 1851
Db 2010 AACCT 2006

RESULT 32
AX777990
LOCUS 458 bp DNA linear PAT 14-JUL-2003
DEFINITION Sequence 147 from Patent WO03039443.
ACCESSION AX777990
VERSION AX777990.1 GI:32694984
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
JOURNAL Hominiidae; Homo.
1
Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
Dugas,M., Ellis,R., Brors,B. and Mergenthaler,S.
Novel genetic markers for leukemias
Patent: WO 03039443-A 147 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE);
Ludwig-Maximilian-Universitaet Muenchen (DE);
FD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
FEATURES
source
location/Qualifiers
1..458
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 22.1%; Score 458; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 3.5e-110;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1609 AGAAGAGGCTCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAAGCAAAACCATTGC 1668
Db 1 AGAAGAGGCTCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAAGCAAAACCATTGC 60
Qy 1669 GCAGGAAGAGGACCTGGAAACCATGACCATCCTCAGGCTCAGATGGAAGTTTACTGTTC 1728
Db 61 GCAGGAAGAGGACCTGGAAACCATGACCATCCTCAGGCTCAGATGGAAGTTTACTGTTC 120
Qy 1729 TGATTTTTCATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGGAGCAACTGGC 1788
Db 121 TGATTTTTCATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGGAGCAACTGGC 180
Qy 1789 ATTGAGCTGGCAGTTCTCTGAAAGAGAAATGATGCTTTCGAAGACGGAGCGAGGAGTC 1848
Db 181 ATTGAGCTGGCAGTTCTCTGAAAGAGAAATGATGCTTTCGAAGACGGAGCGAGGAGTC 240
Qy 1849 CTTGATGGAGATGCAGAGTCGTCATGGGGGGAGAACAACTGACTCTGACGAGAGGCTTA 1908
Db 241 CTTGATGGAGATGCAGAGTCGTCATGGGGGGAGAACAACTGACTCTGACGAGAGGCTTA 300
Qy 1909 CTTTGTTCAAAGAGGAGCTGAGACAGAGGACTGGCGGCAACAGCGGAAATATTCGGATTCA 1968
Db 301 CTTTGTTCAAAGAGGAGCTGAGACAGAGGACTGGCGGCAACAGCGGAAATATTCGGATTCA 360
Qy 1969 TTCCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATAGACAGTTACAGATTACCGTAT 2028
Db 361 TTCCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATAGACAGTTACAGATTACCGTAT 420
```

```

QY      2029  GGATTGCATCATTTAAGTCTTGAATGATATACCTCCCCA 2066
      |||||||
Db      421  GGATTGCATCATTTAAGTCTTGAATGATATACCTCCCCA 458

RESULT 33
AF049614
LOCUS      Homo sapiens huntingtin interacting protein HYPL mRNA, partial cds.
DEFINITION
ACCESSION AF049614
VERSION    AF049614.1  GI:3329430
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens

REFERENCE
AUTHORS    Faber, P.W., Barnes, G.T., Srinidhi, J., Chen, J., Gusella, J.F. and
            MacDonald, M.E.
TITLE      Huntingtin interacts with a family of WW domain proteins
JOURNAL    Hum. Mol. Genet. 7 (9), 1463-1474 (1998)
PUBMED     9700202
AUTHORS    Faber, P.W., Barnes, G.T., Srinidhi, J., Chen, J., Gusella, J.F. and
            MacDonald, M.E.
TITLE      Direct Submission
JOURNAL    Submitted (23-FEB-1998) Molecular Neurogenetics, Mass. Gen. Hosp.,
            Bldg 149, 13th Street, Charlestown, MA 02129, USA
FEATURES   source
            1..435
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="testis"
            <1..>435
            /codon_start=3
            /product="huntingtin interacting protein HYPL"
            /protein_id="AAC26850.1"
            /db_xref="GI:3329431"
            /translation="RKESKVDRAVLKSELKLEAEKVLASKQLQMDKMTIAKOE
            ELETWTLRAQMEYVCSDFHAERAAREKIHEKEQLALQVLAKENDAFEDGGKRS
            LMEMQSRHGARTSDSDQQAYLVQRAEDRDWEQRNPIHSC"
CDS
ORIGIN
Query Match      20.9%; Score 433.4; DB 5; Length 435;
Best Local Similarity 99.8%; Pred. No. 1.2e-103;
Matches 434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1542  CAAGAAAAGAGTCAGAAAAGTGGACAGGGCAGTGTGAAGAACTGAGTGAATAAAGTGG 1601
      |||||||
Db      1    CAAGAAAAGAGTCAGAAAAGTGGACAGGGCAGTGTGAAGAACTGAGTGAATAAAGTGG 60

QY      1602  AACTGCGACGAGAGGCTCGGCTTCCAAACAGCTGCAATGATGATGAAATGAAGCAACCA 1661
      |||||||
Db      61    AACTGCGACGAGAGGCTCGGCTTCCAAACAGCTGCAATGATGATGAAATGAAGCAACCA 120

QY      1662  TTGCCAAGCAGGAAGAGGACCTGGAACCATGACCATCTCAGGGCTCAGATGGAAGTTT 1721
      |||||||
Db      121  TTGCCAAGCAGGAAGAGGACCTGGAACCATGACCATCTCAGGGCTCAGATGGAAGTTT 180

QY      1722  ACTGTTCTGATTTTCATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGAGGAAAAGGAGC 1781
      |||||||
Db      181  ACTGTTCTGATTTTCATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGAGGAAAAGGAGC 240

QY      1782  AACTGCGATGCGAGCTGGCAGTCTCTGTAAGAGAGATGATGCTTTTCGAGAGCGGAGGCA 1841
      |||||||
Db      241  AACTGCGATGCGAGCTGGCAGTCTCTGTAAGAGAGATGATGCTTTTCGAGAGCGGAGGCA 300

QY      1842  GGCAGTCCCTTGATGGAGATGCGAGTTCGTATGGGGCGAGAACCAAGTGAATCTCTGACGAG 1901
      |||||||
Db      301  GGCAGTCCCTTGATGGAGATGCGAGTTCGTATGGGGCGAGAACCAAGTGAATCTCTGACGAG 360

```

```

QY      1902  AGGCTTACCTTGTTCAAAGAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTC 1961
      |||||||
Db      361  AGGCTTACCTTGTTCAAAGAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTC 420

QY      1962  CGATTTCATTCCTGCC 1976
      |||||||
Db      421  CGATTTCATTCCTGCC 435

RESULT 34
AX969999
LOCUS      Sequence 802 from Patent EP1104808.
DEFINITION
ACCESSION AX969999
VERSION    AX969999.1  GI:40976917
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens

REFERENCE
AUTHORS    Dumas Milne Edwards, J.B., Jobert, S. and Giordano, J.Y.
TITLE      ESTs and encoded human proteins
JOURNAL    Patent: EP 1104808-A 802 06-JUN-2001;
            Genset (FR)
FEATURES   Location/Qualifiers
            source
            1..444
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            191..>442
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAF15107.1"
            /db_xref="GI:40976918"
            /translation="MKKRLQKCKQALERKNSAI PSELNEKQLVVTNKKLEQVPSML
            SEIRMEQAKTEDEKSKVLQWTHNKKLQEHNNALTKIE"
ORIGIN
Query Match      20.4%; Score 424.6; DB 2; Length 444;
Best Local Similarity 98.0%; Pred. No. 2.7e-101;
Matches 438; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY      1093  TTCAGATTTTGAAAAAACAAGTAATCGTTTCTGAGATTGAAACCCAGACAGAGGGGAG 1152
      |||||||
Db      1    TTCAGATTTTGAAAAAACAAGTAATCGTTTCTGAGATTGAAACCCAGACAGAGGGGAG 60

QY      1153  CACAGAGAAAGAGATGATGAGAGAAAGGCCCGGAGACTGTGGAGCGNAAGTGAAGC 1212
      |||||||
Db      61    CACAGAGAAAGAGATGATGAGAGAAAGGCCCGGAGACTGTGGAGCGNAAGTGAAGC 120

QY      1213  ACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAAACCTCAG 1272
      |||||||
Db      121  ACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAAACCTCAG 180

QY      1273  CGAAGCTGAGCTTAATCAAGAAAGAGACTTCAAGAAAAGTGTCAAGGCCCTTGAAAGGAAAA 1332
      |||::|
Db      181  CGAMBDCATA--AATGAAGAGAGACTTCAAGAAAAGTGTCAAGGCCCTTGAAAGGAAAAA 237

QY      1333  TTCTGCAATTCATCAGAGTTGAATGAAGAACAGAGCTTGTTTATCTAACAACAAAGTT 1392
      |||||||
Db      238  TTCTGCAATTCATCAGAGTTGAATGAAGAACAGAGCTTGTTTATCTAACAACAAAGTT 297

QY      1393  AGAGCTTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGGA 1452
      |||||||
Db      298  AGAGCTTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGGA 357

QY      1453  TGAAGAAGTCCAAATTAATCTGCTCAGATGACACACAACAAGCTTCTTCAAGAAACATAA 1512
      |||||||
Db      358  TGAAGAAGTCCAAATTAATCTGCTCAGATGACACACAACAAGCTTCTTCAAGAAACATAA 417

```


QY 1513 TAATGCATTGAAAAACAATTGAGGAAGT 1539
|||||
Db 418 TAATGCATTGAAAAACAATTGAGGAAGT 444
|||||
RESULT 35
BD108718
LOCUS BD108718 444 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD108718
VERSION BD108718.1 GI:23203536
KEYWORDS JP 2002010789-A/795.
SOURCE Homo sapiens (human).
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 444)
Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
EST and encoded human protein
Patent: JP 2002010789-A 795 15-JAN-2002;
GENSET CORP
COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/795
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT CDS 191..442.
FEATURES
source
1..444
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 20.4%; Score 424.6; DB 2; Length 444;
Best Local Similarity 98.0%; Pred. No. 2.7e-101;
Matches 438; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
QY 1093 TTCAGATTTTGAAGAAGAAAAACAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAG 1152
|||||
Db 1 TTCAGATTTTGAAGAAGAAAAACAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAG 60
|||||
QY 1153 CACAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTGGAGCGAAGTGGAGC 1212
|||||
Db 61 CACAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTGGAGCGAAGTGGAGC 120
|||||
QY 1213 ACTGAACCTCCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAG 1272
|||||
Db 121 ACTGAACCTCCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAG 180
|||||
QY 1273 CGAAGCTGAGCTTAATGAAGAAGAGACTTCAAGAAAAAGTGTCAAGCCCTTGAAGGAAAAA 1332
|||||
Db 181 CGAMBDCATA---AATGAAGAAGAGACTTCAAGAAAAAGTGTCAAGCCCTTGAAGGAAAAA 237
|||||
QY 1333 TTCTGCAATTCCATCAGAGTTGAATGAAAGCAAGAGCTTGTATATACTAAACAAAAGTT 1392
|||||
Db 238 TTCTGCAATTCCATCAGAGTTGAATGAAAGCAAGAGCTTGTATATACTAAACAAAAGTT 297
|||||
QY 1393 AGAGCTACAAGTGGAAAGCATGCTATCAGAAAATCAAAATGGAACAGGCTAAAAACAGAGGA 1452
Db 298 AGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGGA 357
QY 1453 TGAAGAAGTCCAATTAACTGTGCTACAGATGACACACAAACAGCTTCTTCAAGAACATAA 1512
Db 358 TGAAGAAGTCCAATTAACTGTGCTACAGATGACACACAAACAGCTTCTTCAAGAACATAA 417
|||||
QY 1513 TAATGCATTGAAAAACAATTGAGGAAGT 1539
|||||
Db 418 TAATGCATTGAAAAACAATTGAGGAAGT 444
|||||
RESULT 37
CQ923338
LOCUS CQ923338 493 bp DNA linear PAT 23-NOV-2004
DEFINITION Sequence 4538 from Patent WO2004097052.
ACCESSION CQ923338
KEYWORDS CQ923338.1 GI:56213279

Db 358 TGAAGAAGTCCAATTAACTGTGCTACAGATGACACACAAACAGCTTCTTCAAGAACATAA 417
|||||
QY 1513 TAATGCATTGAAAAACAATTGAGGAAGT 1539
|||||
Db 418 TAATGCATTGAAAAACAATTGAGGAAGT 444
|||||
RESULT 36
AR413165
LOCUS AR413165 444 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 802 from patent US 6639063.
ACCESSION AR413165
VERSION AR413165.1 GI:40168275
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 444)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 802 28-OCT-2003;
Genset S.A.;;
WOX;
FEATURES
source
1..444
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 20.4%; Score 424.6; DB 2; Length 444;
Best Local Similarity 98.0%; Pred. No. 2.7e-101;
Matches 438; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
QY 1093 TTCAGATTTTGAAGAAGAAAAACAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAG 1152
|||||
Db 1 TTCAGATTTTGAAGAAGAAAAACAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAG 60
|||||
QY 1153 CACAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTGGAGCGAAGTGGAGC 1212
|||||
Db 61 CACAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTGGAGCGAAGTGGAGC 120
|||||
QY 1213 ACTGAACCTCCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAG 1272
|||||
Db 121 ACTGAACCTCCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAG 180
|||||
QY 1273 CGAAGCTGAGCTTAATGAAGAAGAGACTTCAAGAAAAAGTGTCAAGCCCTTGAAGGAAAAA 1332
|||||
Db 181 CGAMBDCATA---AATGAAGAAGAGACTTCAAGAAAAAGTGTCAAGCCCTTGAAGGAAAAA 237
|||||
QY 1333 TTCTGCAATTCCATCAGAGTTGAATGAAAGCAAGAGCTTGTATATACTAAACAAAAGTT 1392
|||||
Db 238 TTCTGCAATTCCATCAGAGTTGAATGAAAGCAAGAGCTTGTATATACTAAACAAAAGTT 297
|||||
QY 1393 AGAGCTACAAGTGGAAAGCATGCTATCAGAAAATCAAAATGGAACAGGCTAAAAACAGAGGA 1452
Db 298 AGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGGA 357
QY 1453 TGAAGAAGTCCAATTAACTGTGCTACAGATGACACACAAACAGCTTCTTCAAGAACATAA 1512
Db 358 TGAAGAAGTCCAATTAACTGTGCTACAGATGACACACAAACAGCTTCTTCAAGAACATAA 417
|||||
QY 1513 TAATGCATTGAAAAACAATTGAGGAAGT 1539
|||||
Db 418 TAATGCATTGAAAAACAATTGAGGAAGT 444
|||||
RESULT 37
CQ923338
LOCUS CQ923338 493 bp DNA linear PAT 23-NOV-2004
DEFINITION Sequence 4538 from Patent WO2004097052.
ACCESSION CQ923338
KEYWORDS CQ923338.1 GI:56213279

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1
AUTHORS     Burczynski, M.E., Twine, N.C., Slonim, D.K., Trepicchio, W.L.,
            Straßh, A., Immerman, F. and Dörner, A.J.
TITLE       Methods for prognosis and treatment of solid tumors
JOURNAL     Patent: WO 2004/097052-A 4538 11-NOV-2004;
            Wyeth (US); Burczynski, Michael E. (US)
FEATURES   Location/Qualifiers
            source
            1..493
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            misc_feature
            37..73
            /note="n includes a, c, g, or t, or contains no
            nucleotide"
            misc_feature
            104
            /note="n includes a, c, g, or t, or contains no
            nucleotide"
            misc_feature
            106..166
            /note="n includes a, c, g, or t, or contains no
            nucleotide"
            misc_feature
            193..242
            /note="n includes a, c, g, or t, or contains no
            nucleotide"
ORIGIN
Query Match      15.0%; Score 312; DB 2; Length 493;
Best Local Similarity 69.2%; Pred. No. 2.5e-71;
Matches 334; Conservative 0; Mismatches 147; Indels 2; Gaps 2;
QY 1597 ACTGGAACCTGGCAGAGAGAGGCTCTGGCTTCCAAACA-GCTGCAAAATGGATGAATGAAGC 1655
            |||||
            1 ACTGGAACCTGGCAGAGAGAGGCTCTGGCTTCCAAACANNNNNNNNNNNNNNNNNNN 60
QY 1656 AAACATTCGCCAAGCAGAGAGAGACCTGGAAACCATGACCAT-CCTCAGGAGGCTCAGATG 1714
            |||||
            61 NNNNNNNNNNNNGCAGGAAGAGGACCTGGAAACCATGACCATCNCNNNNNNNNNNNN 120
QY 1715 GAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCAGCAGAGAGAGAAATTCATGAGGAA 1774
            |||||
            121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180
QY 1775 AAGGAGCAACTGGCATTGCGAGCTGGCAGTTCTGCTGAAAGAGAGATGATGCTTTTCGAAGAC 1834
            |||||
            181 AAGGAGCAACTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY 1835 GGAGGAGGAGCTCCTTGTATGGAGATGCGAGTCTGTCATGGGGCGAGAACAAAGTACTCT 1894
            |||||
            241 NNAGGAGGAGCTCCTTGTATGGAGATGCGAGTCTGTCATGGGGCGAGAACAAAGTACTCT 300
QY 1895 GACGAGGAGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGG 1954
            |||||
            301 GACGAGGAGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGG 360
QY 1955 AATATTCGATTCATTCCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATAGACAGTTA 2014
            |||||
            361 AATATTCGATTCATTCCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATAGACAGTTA 420
QY 2015 CAGATTCAGCTGATGATTCATTCATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 2074
            |||||
            421 CAGATTCAGCTGATGATTCATTCATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 480
QY 2075 GGT 2077
            |||
            481 GGT 483

RESULT 39
LOCUS      BC088724
DEFINITION Xenopus laevis hypothetical LOC496250, mRNA (cdna clone MGC:99383
IMAGE:7205613), complete cds.
ACCESSION BC088724
VERSION   BC088724.1
KEYWORDS  MGC.

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1
AUTHORS     Burczynski, M.E., Twine, N.C., Slonim, D.K., Trepicchio, W.L.,
            Straßh, A., Immerman, F. and Dörner, A.J.
TITLE       Methods for prognosis and treatment of solid tumors
JOURNAL     Patent: WO 2004/097052-A 4538 11-NOV-2004;
            Wyeth (US); Burczynski, Michael E. (US)
FEATURES   Location/Qualifiers
            source
            1..493
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            misc_feature
            37..73
            /note="n includes a, c, g, or t, or contains no
            nucleotide"
            misc_feature
            104
            /note="n includes a, c, g, or t, or contains no
            nucleotide"
            misc_feature
            106..166
            /note="n includes a, c, g, or t, or contains no
            nucleotide"
            misc_feature
            193..242
            /note="n includes a, c, g, or t, or contains no
            nucleotide"
ORIGIN
Query Match      15.0%; Score 312; DB 2; Length 493;
Best Local Similarity 69.2%; Pred. No. 2.5e-71;
Matches 334; Conservative 0; Mismatches 147; Indels 2; Gaps 2;
QY 1597 ACTGGAACCTGGCAGAGAGAGGCTCTGGCTTCCAAACA-GCTGCAAAATGGATGAATGAAGC 1655
            |||||
            1 ACTGGAACCTGGCAGAGAGAGGCTCTGGCTTCCAAACANNNNNNNNNNNNNNNNNNN 60
QY 1656 AAACATTCGCCAAGCAGAGAGAGACCTGGAAACCATGACCAT-CCTCAGGAGGCTCAGATG 1714
            |||||
            61 NNNNNNNNNNNNGCAGGAAGAGGACCTGGAAACCATGACCATCNCNNNNNNNNNNNN 120
QY 1715 GAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCAGCAGAGAGAGAAATTCATGAGGAA 1774
            |||||
            121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180
QY 1775 AAGGAGCAACTGGCATTGCGAGCTGGCAGTTCTGCTGAAAGAGAGATGATGCTTTTCGAAGAC 1834
            |||||
            181 AAGGAGCAACTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY 1835 GGAGGAGGAGCTCCTTGTATGGAGATGCGAGTCTGTCATGGGGCGAGAACAAAGTACTCT 1894
            |||||
            241 NNAGGAGGAGCTCCTTGTATGGAGATGCGAGTCTGTCATGGGGCGAGAACAAAGTACTCT 300
QY 1895 GACGAGGAGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGG 1954
            |||||
            301 GACGAGGAGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGG 360
QY 1955 AATATTCGATTCATTCCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATAGACAGTTA 2014
            |||||
            361 AATATTCGATTCATTCCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATAGACAGTTA 420
QY 2015 CAGATTCAGCTGATGATTCATTCATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2074
            |||||
            421 CAGATTCAGCTGATGATTCATTCATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 480
QY 2075 GGT 2077
            |||
            481 GGT 483

RESULT 38
CS135191
```


1080	ACAGAGCGCTGAAAAAATCAACTAGGCATCTTGCAGGTTTCATATGATAAATTAACCTCA	1139
1505	GAACATATAATGCAATTCAGGAAACAATTCAGGCAATCAACAGAAAAAGAGTCAGAAAAAGTG	1564
1140	GAATATC-----AGGAACCTAAGGATTAGAGAGATTGAAAAAGTT	1178
1565	GACAGGCGAGTCGTGAAGGAACCTGAGTGAATAAACTGGAACTGGCAGAGAAAGGCTCTGGCT	1624
1179	TCCAAAGTTGAGTTTATGAGCTGCTTGAAGAGCTGGATGATGTGTAAGAAAGGCACATTGCA	1238
1625	TCCAAACAGCTGCAAATGGATGAAATGAAGCAAAACCAATTCGCCAAGCAGGAAGAGGACCTG	1684
1239	AAAAACCAATTTGAAATTTGATGAAATGAGAGAATGGAACAAAAACAGAGGAGGATAAA	1298
1685	GAAACCATGACCATCTCTCAGGGCTCAGATGGAAAGTTTACTGTCTGATTTTCATGCTGAA	1744
1299	GAACCAATTTGAACTTCTTCGTGCTCAGGTTTGAACGTGTACTGTGCAGATTTCCATGCTGAA	1358
1745	AGAGCAGCGAGAGAGAAATTCATGAGGAAAAAGAGCAACTGGCCATTGCGAGCTGGCGCATTT	1804
1359	AGATCAGCAGAGAAACATTCATCAGAGAGAGGAGCAGTTAGCAACTCGCTCGGCATAT	1418
1805	CTGCTGAAAGAGAGAAATGATGCTTTGGAACGGA-----GGCAGGCAGTCCCTGATGGAG	1858
1419	ATGATACAAGAAATATGAAAACTGAAGGAAGAAATGATGGGCAAGCAGTCAATAGAGCAG	1478
1859	ATGCAGAGTCGTCAATGGGCGAGAAA---CAAGTCACTCTGACCAAGGCTTACCTTGTT	1915
1479	CTCAGAGGCGCCATGTTGTCACACAGCTTATTGGATCCAGTGAAGGCTCTTATTATG	1538
1916	CAAGAGGAGCTGAGGACAGGCACTGGCGGCAACAGCGGAATATTCGATTCATTCCTGTC	1975
1539	GCAAGAGGGGAGCAACATGG-----AGCAGGCAAGTATTACTGTGTATACATGT	1589
1976	CCCAAGTGTGGAGAGTTCTTCCTGTGACATAGACACGTTTACAGATTACGCTGATGGATTGC	2035
1590	CCAAATGTAACTGACTGTTTCCAGACATGGATACGCTGCAGATTCACGTCATGGAATGT	1649
2036	ATCATTTTAAAG 2045	
1650	ATCACCTTAAAG 1659	

RESULT	40
CS213387	
LOCUS	505 bp DNA linear PAT 15-DEC-2005
DEFINITION	Sequence 2706 from Patent WO2005111246.
ACCESSION	CS213387
VERSION	CS213387.1 GI:83697404
KEYWORDS	.
SOURCE	Cricetulus griseus (Chinese hamster)
ORGANISM	Cricetulus griseus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Cricetidae; Cricetinae; Cricetus.
REFERENCE	1 Melville,M.W
AUTHORS	Oligonucleotide arrays to monitor gene expression and methods for making and using same
TITLE	Patent: WO 2005111246-A 2706 24-NOV-2005;
JOURNAL	Wyeth (US)
FEATURES	Location/Qualifiers
source	1..505 /organism="Cricetulus griseus" /mol_type="unassigned DNA" /db_xref="taxon:10029"

289	Qy	CTTTTCCACAGGAACTTTCGCAATGTCCATCAACCTCTCAGCTGCCTCACTGAAAGGA	348
83	Db	CTAAGCGGACGGACCTGTTACCATGTCCCATCAACCTCCGAGCTGCTGACTGAGAAGGG	142
349	Qy	GGACAGCCCCAGTGAAGACACAGAAATGACCCCCCCACCTGGCCGCCACCCAAACCTGGA	408
143	Db	TGACAGGCTCTGTGAGACCCCGAATAATGNAACCCCTTAATAGGTTTCACTCCAGCCTGGA	202
409	Qy	CACGTTTACCCCCGGAGAGCTGTGCAGCAGATGAAAGAGCTCCTGTGCCAAGAACCACCA	468
203	Db	CACATTCACTCTCTGAGGAGATGCTGCAGCAAAATGAAAGGAATCTCGTGTGAGAACCACCA	262
469	Qy	GCTGAAAGAGCCATGAAGCTTAATTAATCAAGCCATCAAGGGAGATTTTCAGGAGCTTTC	528
263	Db	GCTGAAAGAGCCATGAAGCTTAAACCAATCAAGCCATGAAGGGAGATATGAGGAGCTTTC	322
529	Qy	GGCTCTGACAGAGAAACAGAAAGGAAGAACCCAGTTTTTTTGAGATACAGAGCAAAAGAAGC	588
323	Db	AGCTGGGCAGAGAACGAGAAGGAAGAACACAGTGTGTTTGACACACAGAAACAAAGAAGC	382
589	Qy	AAAAGAGGCTCTAATGGCCCTTGAGTCAATGAGAAATGAGAAATTTGAAGGAAGAGCTTGAAA	648
383	Db	TAAAGAGCTCTGATGACCTTTGAGTTATGAAATGAGAAACTGAAGGTGAGGCTTGAAA	442
649	Qy	ACTTAAAGGGAAATCAGAAAGGTCATCTGAGGACCCCACTGATGACTCCAGGGTTCCTCCAG	708
443	Db	ACTTAAGAGAGAGTCAGAAAGGCCATTTGAAGGCTCCATATGTGGCTCCAGGCTCCCAT	502
709	Qy	GG 710	
503	Db	GG 504	

RESULT 41	
CS217029	CS217029
LOCUS	Sequence 6348 from Patent WO2005111246.
DEFINITION	505 bp DNA
ACCESSION	CS217029
VERSION	CS217029.1 GI:83691850
KEYWORDS	
SOURCE	Cricetulus griseus (Chinese hamster)
ORGANISM	Cricetulus griseus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
	Sciurognath; Muroidae; Cricetidae; Cricetinae; Cricetulus.
REFERENCE	1 Melville,M.W.
AUTHORS	Oligonucleotide arrays to monitor gene expression and methods for
TITLE	making and using same
JOURNAL	Patent: WO 2005111246-A 6348 24-NOV-2005;
	Wyeth (US)
FEATURES	Location/Qualifiers
source	1..505
	/organism="Cricetulus griseus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:10029"

Query Match	Score	DB 2	Length
Best Local Similarity	80.8%	292.4	505
Matches 341	Conservative	Pred. No. 4e-66	
		0; Mismatches	81; Indels
		0; Gaps	0;

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
 Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
 Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lissa Prabh
 Parvaneh Saeedi, JR Santos, Angeli Schnerch, Ursula Skalska,
 Duane Smalish, Jeff Stott, Miranda Tsai, George Yang, Jacque
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 147 Row: O Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 39645451.

FEATURES

source

1..2536
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="MGC:77868 IMAGE:7002497"
 /tissue_type="whole, adult zebrafish"
 /clone_lib="NIH_ZGC_10"
 /lab_host="DH10B"
 /note="Vector: pExpress1"
 /note="pExpress1"
 1..2536
 /gene="zgc:77868"
 /note="synonyms: fj52f04, OPTN, MGC66386, zgc:66386"
 /db_xref="GeneID:336159"
 /db_xref="ZFIN:ZDB-GENE-030131-8103"
 233..1786
 /gene="zgc:77868"
 /codon_start=1
 /product="optineurin"
 /protein_id="AAH63986.1"
 /db_xref="GI:39645452"
 /db_xref="GeneID:336159"
 /db_xref="ZFIN:ZDB-GENE-030131-8103"
 /translation="MNGDISHPRGSGNGLSLETLQOMNTLIKENRDLKALQNT
 LSKRPFGLSAWKQKQERDFLEORLEAETRLNMDVENEALKNQVKELEKSGAE
 CLHTEALRGQLRIQAEKNDLVANSLQMGOGSPNSFIRIATDDLLKVTKD
 LSSVPEASFPKAESEETVRLSLRAETDEKRLQLTQEARGRIAELESKLE
 HADSSAQTSLPSAETNASTEVNLEDLILCNELKQAIKLEAESMKRNLDRCK
 DLEODLTGLTOLGDKQVQAEKNDCLKVOMESLQAAIKLEQKKTODEKNLNLQDAY
 TKLFEDYSLEKKKRESKSVSKDYDELQTFATAEKALADKQKIDEMKMLFOKE
 KOLETISVFOAQAEIYSSDFYAEARAAREKIHKEERLATQLEYVKKQSLQSEMESL
 GRHSMSEMRHVRPGANPQGTAPNNLPGRGEWQQQNIIPDHACPKCGEVLPLDLSL
 QIHIMDCIT"

gene

CDS

1190 ACTGTTGGAAGCAAGTGAAGCACTGAACCTCCAGTGCATCTCTGTTTAAGAGAGCTT 1249
 935 AATGCTAGCACTGAGGTGAAGATCTGGAAGATCAGTTGTTGAGCTCTGTAATGAGTTG 994
 1250 CAAGAGGCTCATACAAAATCTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAG 1309
 995 AAGCAAGCTCAGATTAACTGGACGAGGCTGAGAGCATGAAAGAAACCTCGCAGGACAGA 1054
 1310 TGTCAGGCTTGAAGGAAAATCTGCAATCCATCAGCTTGAATGAAGAGCAAGAG 1369
 1055 TGTAAAGATTGAGCAGGATTGGGCACTTTAAAGACTCAGCTCGGAGACAAACAGAG 1114
 1370 CTTGTTTATTAACAAAAGTTAGAGCTACAAGTGAAGAGCATGCTATCAGAAAATCAAA 1429
 1115 GTCCAGGCTGAGAACGACTCTCTGAAGGTTTCAGATGTAAGAGCTCCAAAGCGGCATCAA 1174

ORIGIN

Query Match 9.4%; Score 195.4; DB 11; Length 2536;
 Best Local Similarity 55.2%; Pred. No. 1.8e-40;
 Matches 479; Conservative 0; Mismatches 362; Indels 27; Gaps 4;

QY 1430 ATGGAACAGGCTAAAAACAGAGAGTGAAGAGTCCAAATTAACATGCTGTACAGATGACACAC 1489
 DB 1175 CTGGAACAGAGAAAACCCAGATGAGGAACAAATCTAAACCAGCTGAAGATCGGTAC 1234
 QY 1490 AACAGCTTCTTCAAGAACATAATAATGATGAAAAACAATTGAGGAATTAACAAGAAAA 1549
 DB 1235 ACCAAGCTGTTTGAAGACTACAGTGAGCTCCAGGAGGAA-----AAGAAAGAAGA 1283
 QY 1550 GAGTCAGAAAAGTGGACAGGCGAGTGTGAAGNACTGAGTGAAGAACTGGAATGGAAC 1609
 DB 1284 GAGAGAGTTGTATCAAGAGCAGATTATGATGAATCTCAACACCAAGGTTTGAACCT-GCT 1342
 QY 1610 GAGAGGCTCTGGCTTCCAAACAGCTGCAAAATGGAATGAATGAACCAACCAATGCGCAAG 1669
 DB 1343 GAAAGGCACTTTCGATATAACAGCAGAAAATTCGATGAGATGAAGATGAAGTATTCAG 1402
 QY 1670 CAGGAAGAGGACCTGGAAACCATGACCATCTCCAGGCTCAGATGGAAGTTTACTGTTCT 1729
 DB 1403 AAGGAAGAGATCTGGAACCACTTCTGTTTCCAGGCACAGGCAGAAATCTATTCTCTCA 1462
 QY 1730 GATTTCATGCTGAAAGAGCAGCAGAGAGAAAATTCATGAGGAAAAGAGCAACTGGCA 1789
 DB 1463 GATTTCTACGCAAGAAAGAGCAGCTCGGAGAAGATTTCAGAGGAGAAAGAGCGTTTGGCC 1522
 QY 1790 TTGCAGCTGGCAGTTCTGCTGAAAAGAGA-----ATGATGCTTTCCGAAGACGGA 1837
 DB 1523 ACTCAGCTGGAGTATGTGAAAAGCAGAACAGCCAGCTTCAAGAGAGATGAGTCACTT 1582
 QY 1838 GCGAGCAGTCTCTGATGAGATGAGAGTCTCATGGGGCGGAGAACAAAGTCACTCTGAC 1897
 DB 1593 GGGAGCATCAATGAGCGAGATGCAAGAGAGCATGTGCC--ACGAGGAGCAATCCT 1639
 QY 1898 CAGCAGGCTTACCTTGTTCAAAGAGAGCTGAGGACAGGAGTGGCGCAACAGCGGAT 1957
 DB 1640 CAAGGGCCACAGCTCCAAACAAACCTGCTGGAGTAGAGGTGAATGGCAGCAGCAGAG 1699
 QY 1958 ATTCGATTCATCTCTGCCCAAGTGTGAGAGGTTCTGCTGACATAGACAGCTTACAG 2017
 DB 1700 ATTCAGATCAGCTGTGCTAAATGTGGGAAGTCTACCTGACTGACTGCTCCCTGCGAG 1759
 QY 2018 ATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2045
 DB 1760 ATTCACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1787
 RESULT 47
 AY749110S02
 LOCUS AY749110S02 257 bp DNA linear PRI 27-JUN-2005
 DEFINITION Macaca mulatta optineurin (OPTN) gene, exon 2.
 ACCESSION AY749111
 VERSION AY749111.1 GI:54111692
 KEYWORDS
 SEGMENTS
 SOURCE 2 of 13
 ORGANISM Macaca mulatta (rhesus monkey)
 Macaca mulatta
 Macaca mulatta
 REFERENCE 1 (bases 1 to 257)
 AUTHORS Rezaie,T., Waitzman,D.M., Seeman,J.L., Kaufman,P.L. and Sarfarazi,M.
 TITLE Molecular cloning and expression profiling of optineurin in the rhesus monkey
 JOURNAL Invest. Ophthalmol. Vis. Sci. 46 (7), 2404-2410 (2005)
 PUBLISHED 15980228
 REFERENCE 2 (bases 1 to 257)
 AUTHORS Rezaie,T. and Sarfarazi,M.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-2004) Molecular Ophthalmic Genetics Laboratory, Surgical Research Center, Department of Surgery, University of Connecticut Health Center, 263 Farmington Ave., Farmington, CT 06030-1110, USA
 FEATURES Location/Qualifiers

```
source 1..257
/organism="Macaca mulatta"
/mol_type="genomic DNA"
/db_xref="taxon:9544"
33..235
/gene="OPTN"
/number=2

ORIGIN
Query Match 9.3%; Score 193.2; DB 5; Length 257;
Best Local Similarity 96.1%; Pred. No. 1.2e-39;
Matches 198; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 475 AGAGCCATGAGCTAATAATCAAGCATGAAGGAGATTGAGGAGCTTTCGGCTG 534
Db |||
31 AGAAGCCATGAGCTAATAATCAAGCATGAAGGAGCTTTCGGCTG 90
Qy 535 GACAGAGAAACAGAGGAGAACCCAGTCTTTTGTAGATCAGAGCAAGNAGCAAGA 594
Db |||
91 GACAGAGAAACAGAGGAGAACCCAGTCTTTTGTAGATCAGAGCAAGNAGCAAGA 150
Qy 595 GCGTCTAATGCGCTTGAGTCATGAGATGAGAAATGAAGGAGAGCTTGGAAACTAAA 654
Db |||
151 GCGTCTAATGCGCTTGAGTCATGAGATGAGAAATGAAGGAGAGCTTGGAAACTGAA 210
Qy 655 AGGAAATCAGAAAGGTCATCTGAGG 680
Db |||
211 AGGAAATCAGAAAGGTCATCTGAGG 236

RESULT 48
CQ923519
LOCUS 193 bp DNA linear PAT 23-NOV-2004
DEFINITION Sequence 4719 from Patent WO2004097052.
ACCESSION CQ923519
VERSION CQ923519.1 GI:56213460
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Burczynski, M.E., Twine, N.C., Slonim, D.K., Trepicchio, W.L.,
Strahs, A., Immerman, F. and Dörner, A.J.
TITLE Methods for prognosis and treatment of solid tumors
JOURNAL Patent: WO 2004097052-A 4719 11-NOV-2004;
Wyeth (US); Burczynski, Michael E. (US)
FEATURES
source
1..193
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 9.3%; Score 193; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.4e-39;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1843 GCAGTCTTGATGAGATGAGAGTCGTCTATGGGGCGAGAACAGTGACTCTGACCA 1902
Db |||
1 GCAGTCTTGATGAGATGAGAGTCGTCTATGGGGCGAGAACAGTGACTCTGACCA 60
Qy 1903 GGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGGAGCTGGCGGCAACAGCGGA 1962
Db |||
61 GGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGGAGCTGGCGGCAACAGCGGA 120
Qy 1963 GATTCAATTCCTGCCCAAGTGTGGAGAGGTTTCCTCGCATAGACATAGACACGTTACAGATTCA 2022
Db |||
121 GATTCAATTCCTGCCCAAGTGTGGAGAGGTTTCCTCGCATAGACATAGACACGTTACAGATTCA 180
Qy 2023 CGTGATGGATTGC 2035
Db |||
181 CGTGATGGATTGC 193

RESULT 50
BC055628
LOCUS 2075 bp mRNA linear VRT 24-NOV-2004
DEFINITION Danio rerio zgc:77868, mRNA (cdna clone MGC:66386 IMAGE:5777373),
complete cds.
ACCESSION BC055628
VERSION BC055628.1 GI:33416920
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 2075)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ussidi, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
```

```
Db 181 CGTGATGGATTGC 193

RESULT 49
CS135190
LOCUS 193 bp DNA linear PAT 09-AUG-2005
DEFINITION Sequence 447 from Patent WO2005068655.
ACCESSION CS135190
VERSION CS135190.1 GI:72065947
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Smith, S.K.
TITLE Methods of assessing a tissue inflammatory response using
expression profiles of endothelial cells
JOURNAL Patent: WO 2005068655-A 447 28-JUL-2005;
Cambridge University Technical Services Limited (GB)
FEATURES
source
1..193
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Probe 41742_s at HG-U95Av2"

ORIGIN
Query Match 9.3%; Score 193; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.4e-39;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1843 GCAGTCTTGATGAGATGAGAGTCGTCTATGGGGCGAGAACAGTGACTCTGACCA 1902
Db |||
1 GCAGTCTTGATGAGATGAGAGTCGTCTATGGGGCGAGAACAGTGACTCTGACCA 60
Qy 1903 GGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGGAGCTGGCGGCAACAGCGGA 1962
Db |||
61 GGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGGAGCTGGCGGCAACAGCGGA 120
Qy 1963 GATTCAATTCCTGCCCAAGTGTGGAGAGGTTTCCTCGCATAGACATAGACACGTTACAGATTCA 2022
Db |||
121 GATTCAATTCCTGCCCAAGTGTGGAGAGGTTTCCTCGCATAGACATAGACACGTTACAGATTCA 180
Qy 2023 CGTGATGGATTGC 2035
Db |||
181 CGTGATGGATTGC 193

RESULT 50
BC055628
LOCUS 2075 bp mRNA linear VRT 24-NOV-2004
DEFINITION Danio rerio zgc:77868, mRNA (cdna clone MGC:66386 IMAGE:5777373),
complete cds.
ACCESSION BC055628
VERSION BC055628.1 GI:33416920
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 2075)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ussidi, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
```

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettner, M., Madan, A., Rodrigues, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2075)
 Director MGC Project.
 Direct Submission
 Submitted (01-AUG-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Dr. Sumio Sugano
 CDNA Library Preparation: Dr. Sumio Sugano
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 122 Row: n Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 33416920.

FEATURES

Location/Qualifiers
 1..2075
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="MGC:66386 IMAGE:577373"
 /tissue_type="whole body, adult male"
 /clone_lib="Sugano SJD adult male"
 /lab_host="DH10B"
 /note="Vector: pME18S-FL3"
 1..2075
 /gene="zgc:77868"
 /note="synonyms: f5j52f04, OPTN, MGC66386, zgc:66386"
 /db_xref="GeneID:336159"
 /db_xref="ZFIN:ZDB-GENE-030131-8103"
 81..1391
 /gene="zgc:77868"
 /codon_start=1
 /product="zgc:77868 protein"
 /protein_id="AAH5628.1"
 /db_xref="GI:33416921"
 /db_xref="GeneID:336159"
 /db_xref="ZFIN:ZDB-GENE-030131-8103"
 /translation="MNGDISHPRGSGFGNIGLSLEETLQMNVLIKENRDLKALQNTN
 LSKERFEGLSAWKEKDFLEQRLEEARLNTMTDVENELKNQVRELEKSGAE
 CLHTELEALQGQILRTQAEKNDLVAMNSQLKMGQGSFNSFIETRIATNASTEVK
 NLEDLKLLKQALQKILDEASMKRNLDQRCDLEODLGLTKLQDGKQVQAEV
 DCLKVOMESLQAAIKLEQKTKDEKSNLNQLDAYTKLFEDYSELQEKKKREGCVSK
 DDYDELQTFATKALADKQKIDEMKMLFKKEDLETISVFQAEIYSSDFVAE
 RAAREKIHKEKRLATQPIYVKQNSQLQEFEMESLGRHSMSEQRHVRPGANPQPT
 APNNLPGGGEWQQQNPIDHACPKGEVLPLDLSUQIHIMDCII"

gene

gene

CDS

CDS

	Matches	477;	Conservative	0;	Mismatches	364;	Indels	27;	Gaps	4;
QY	1190	ACTGTTGAAGCGAAGTGAAGACCTGAACCTTCAGGTGACATCTCTGTTTAAAGAGCTT	1249							
Db	540	AATGCTAGCACTGAGGTGAAGAACTGGAAGATCACTGTTGTTGAAGCTCTGTATGAGTTG	599							
QY	1250	CAAGAGGCTCATACAAAACCTCAGCAAGCTAATGAAGACAGACTTCAAGAAAAAG	1309							
Db	600	AAGCAAGCTCAGATTAACTGACAGGCTGAGACATGAAGAAGAACTCGCAGGACAGA	659							
QY	1310	TGTCAAGGCTTGAAGGAAAAAATCTGCAATTCATCAGAGTTGAATGAAGAAAAAG	1369							
Db	660	TGTAAGATTGGAGCAGGATTTGGGCACTTTAAAGACTCAGCTCGGAGACAAACAGAAG	719							
QY	1370	CTTGTGTTATCTAACAACAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAA	1429							
Db	720	CTCCAGGCTGAGAACGACTGTCTGAAGGTTGAGATGGAAAGCTCCAGCGCGATCAAA	779							
QY	1430	ATGGAAACAGGCTTAAACACAGAGATGAAAGTCCAAATTAACCTGTCTACAGATGACAC	1489							
Db	780	CTGGAAACAGAAAGAAACCCAGATGAGAAAGAGCAATCTAAACCACTGAAAGATGGTAC	839							
QY	1490	AACAGCTTTCTTCAAGAAACATAATATGCAATTTGAAAAACAATTTGAGGAACTAAACA	1549							
Db	840	ACCAAGCTGTTTGAAGACTACAGTCACTCCAGGAGGAA-----AAGAAGAAGA	888							
QY	1550	GAGTCAGAAAAAGTGGACAGGCACTGCTGAAGAACTGAGTGAAGAACTGGAACCTGGCA	1609							
Db	889	GAGAGGGTTGTGTATCAAAAGGACGATTATGATGAATCTCCAAACCAAGTTTGAACCT	947							
QY	1610	GAGAGGCTCTGGCTTCCAAACAGCTGCAAAATGGAATGAATGAAGCAACCACTTGCAC	1669							
Db	948	GAAGAGGCACTTGCAGATAACAGCAGAAATTCGATGAGATGAAGATGAATTCCTCAG	1007							
QY	1670	CAGGAAGAGGACCTCGAAACCATGACCATCTCCAGGCTCAGATGGAAGTTTACTGTTCT	1729							
Db	1008	AGGAAAAAGGATCTGGAACCATTTCTGTTTTCAGGCACAGCAGAAATCTATTCTCTCA	1067							
QY	1730	GATTTTCATGCTGAAAGACGACGAGAGAGAAAAATTCATGAGAGAAAGAGGACACTGGCA	1789							
Db	1068	GATTTCTACGAGAAAGAGCAGCTCGGAGAGAGATTTCAGAGAGAAAGAGCGTTTGGCC	1127							
QY	1790	TTGACGCTGGCAGTTCTGCTGAAAGAGA-----ATGATGCTTTTCAAGACGGA	1837							
Db	1128	ATCAGCTGGAGTATGTGAAGAAAGCAGAACACCCAGCTTCAAGAGGAGATGAGTCACTT	1187							
QY	1838	GGCAGGCACTCTTGTATGAGAGATGAGAGTCTGTCATGGGCGGAGAACAAAGTGACTGAC	1897							
Db	1188	GGGAGGCATTCAATGAGCGAGATGCAAGAGAGGCAATGTGCC--ACGAGGAGCAATCCT	1244							
QY	1898	CAGCAGGCTTACCTTGTTCAGAGAGGCTGAGGACAGGCACTGGCGGCAACAGCGGAAT	1957							
Db	1245	CAAGGGCCCCACAGCTCCAAACCAACCTGCTGAGAGTAGAGGTGAATGGCAGCAGAGAAC	1304							
QY	1958	ATTCGAGTTTCACTCTGCCCAAGTGTGGAGAGGTTTCTGCCCTGACATAGACAGTTTACAG	2017							
Db	1305	ATTCAGATCAGCTTGTCTTAAATGTGGGGAAGTCTTACCTGACCTGACCTCCCTGCAG	1364							
QY	2018	ATTCACGTCAGTGGATTGTCATCTTTAAG	2045							
Db	1365	ATCCACATCATGATTGCATCATCTGAG	1392							

Search completed: May 30, 2006, 03:06:49
 Job time : 11410 secs

ORIGIN

Query Match 9.2%; Score 191.6; DB 11; Length 2075;
 Best Local Similarity 55.0%; Pred. No. 1.9e-39;

c	92	78.6	3.8	444	5	AD138328	Ad138328 Human ova
	93	77.8	3.7	1239	14	AD200649	Ad200649 Mouse NEM
	94	77.8	3.7	1874	10	ADC36537	ADC36537 DNA encod
	95	77.8	3.7	1874	12	ADI61940	ADI61940 Necrosis
	96	77.8	3.7	1874	12	ADI29002	ADI29002 Human NEM
	97	77.8	3.7	1874	13	ADI61943	ADI61943 Necrosis
	98	76.6	3.7	486	13	ADSI17866	ADSI17866 Human iKB
	99	76.6	3.7	1104	13	ADSI17862	ADSI17862 Human iK
	100	76.6	3.7	1260	12	ADK71964	ADK71964 Human tum
	101	76.6	3.7	1967	13	ADQ87415	ADQ87415 Human tum
	102	76.6	3.7	1967	13	ADQ86256	ADQ86256 Human tum
	103	76.6	3.7	1975	8	ADA44712	ADA44712 Human inh
	104	76.6	3.7	1994	3	AAC81426	AAC81426 Human i-K
	105	76.6	3.7	1994	3	AAA35027	AAA35027 Human ade
	106	76.6	3.7	1994	3	AAF21149	AAF21149 Human low
	107	76.6	3.7	1994	8	ADA44705	ADA44705 Human inh
	108	76.6	3.7	1994	10	ABZ96843	ABZ96843 Human nuc
	109	76.6	3.7	1994	10	ABV75393	ABV75393 Human NEM
	110	76.6	3.7	1994	11	ABD20692	ABD20692 Human pul
	111	76.6	3.7	2009	2	AAZ07513	AAZ07513 Human RIP
	112	76.6	3.7	2034	2	AAZ07514	AAZ07514 Human RIP
	113	76.6	3.7	2035	9	ACD23012	ACD23012 Human NEM
	114	76.6	3.7	2035	14	ADZ00655	ADZ00655 Human NEM
	115	76.6	3.7	8631	3	AAA35028	AAA35028 Human ade
	116	76.6	3.7	8631	3	AAF21150	AAF21150 Human low
	117	76.6	3.7	8631	10	ABZ96844	ABZ96844 Human nuc
	118	76.6	3.7	8631	11	ABD20693	ABD20693 Human pul
	119	66.8	3.2	2038	10	ADF82713	ADF82713 Leukaemia
	120	60	2.9	60	6	ABN43434	ABN43434 Human sgl
	121	60	2.9	236	9	ACH15127	ACH15127 Human adu
	122	58.4	2.8	3662	4	ABH05274	ABH05274 Drosophila
	123	54	2.6	2000	8	ADA71938	ADA71938 Rice gene
	124	52.8	2.5	414	8	ABX41363	ABX41363 Bovine ES
	125	52.6	2.5	3579	3	AAA70099	AAA70099 Plasmodiu
	126	51.8	2.5	2517	13	ADU01785	ADU01785 Novel hum
	127	51.2	2.5	627	13	ACN54555	ACN54555 Cotton an
	128	51.2	2.5	1998	3	AAA70212	AAA70212 Plasmodiu
	129	51.2	2.5	6741	3	AAA10595	AAA10595 Gene enco
	130	51	2.5	51	5	ABL00186	ABL00186 Human sil
	131	51	2.5	807	4	AAL21312	AAL21312 Human bre
	132	50.2	2.4	624	13	ACN54604	ACN54604 Cotton an
	133	49.4	2.4	1155	4	AA546749	AA546749 Tumour su
	134	49.2	2.4	446	8	ABX38235	ABX38235 Bovine ES
	135	49.2	2.4	11000	14	AEE04859_0	AEE04859 Cancer-as
	136	48.8	2.3	629	10	ABT21705	ABT21705 Breast ca
	137	48.4	2.3	807	3	AACT75308	AACT75308 Human ORF
	138	48.4	2.3	4262	5	AAF98697	AAF98697 Human ova
	139	48.4	2.3	5142	6	ABS76388	ABS76388 cDNA enco
	140	48.4	2.3	5142	14	ABE35261	ABE35261 Human Gef
	141	48	2.3	344	8	ABX35871	ABX35871 Bovine ES
	142	48	2.3	447	13	ACN54792	ACN54792 Cotton an
	143	47.8	2.3	14006	6	ABL33958	ABL33958 Human imm
	144	47.6	2.3	516	8	ABX40620	ABX40620 Bovine ES
	145	47.4	2.3	5059	2	AA843332	AA843332 Steath v
	146	47.4	2.3	6486	6	ABQ67050	ABQ67050 Human ang
	147	47.4	2.3	6668	6	ABL33697	ABL33697 Human imm
	148	47.2	2.3	2943	3	AAA70229	AAA70229 Plasmodiu
	149	47	2.3	863	4	AAI95036	AAI95036 Human neu
	150	47	2.3	1965	3	AAA70096	AAA70096 Plasmodiu

ALIGNMENTS

RESULT 1
ACF03996
ID ACF03996 standard; cDNA; 2077 BP.
XX ACF03996;
AC ACF03996;
XX
DT 23-SEP-2003 (first entry)
XX
XX Human optineurin isoform 1 encoding cDNA SEQ ID NO:1.
DE

Human; optineurin; OPTN; chromosome 10; 10p14; glaucoma; ophthalmic;
gene therapy; gene; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 311..2044
/tag= a
/product= "optineurin isoform 1"
/replace(412,A)
/tag= b
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(433,A)
/tag= c
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(458,A)
/tag= d
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(603,A)
/tag= e
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(799,G)
/tag= f
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(911,T)
/tag= g
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(947,C)
/tag= h
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(949,C)
/tag= i
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(958,G)
/tag= j
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(1379,C)
/tag= k
variation
/standard_name= "single nucleotide polymorphism (SNP)"
replace(1944,A)
/tag= l
variation
/standard_name= "single nucleotide polymorphism (SNP)"
WO2003056037-A1.
10-JUL-2003.
18-DEC-2002; 2002WO-US041116.
24-DEC-2001; 2001US-0344754P.
30-JAN-2002; 2002US-00060981.
28-FEB-2002; 2002US-00090118.
25-OCT-2002; 2002US-00281454.
(UYCO-) UNIV CONNECTICUT.
(SGEO-) ST GEORGES ENTERPRISES LTD.
Sarfarazi M, Rezaie T, Child AH;
WPI; 2003-598273/56.
P-PSDB; ABR82077.
Diagnosing the presence or absence of optineurin-associated glaucoma or
optineurin-associated increased risk of glaucoma in an individual by
assessing a test sample for the presence or absence of an alteration in
the optineurin gene.
Claim 1; Page 68-71; 83pp; English.
The present invention describes a method (M1) for diagnosing the presence
or absence of optineurin-associated glaucoma or an optineurin-associated
increased risk of glaucoma in an individual comprising detecting the

CC presence or absence of an alteration in expression, composition or
CC activity of an optineurin nucleic acid or polypeptide, which is
CC indicative of the presence or absence, respectively, of the optineurin-
CC associated glaucoma. Optineurin has ophthalmic activity and can be used
CC in gene therapy. Optineurin therapeutic agents can be used for the
CC manufacture of a medicament for the treatment of glaucoma or of an
CC increased risk for glaucoma. Human optineurin is located on chromosome
CC 10, more specifically to 10p14. The present sequence encodes human
CC optineurin isoform 1, which is used in the exemplification of the present
CC invention
XX
SQ

Sequence 2077 BP; 682 A; 431 C; 566 G; 398 T; 0 U; 0 Other;
Query Match 99.9%; Score 2075.4; DB 9; Length 2077;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCCGTCCGGAGTCTCTCCAGCGCGCAGGATGCCAGAAACATGACCTGAGCGA 60
DB 1 ATCCCGTCCGGAGTCTCTCCAGCGCGCAGGATGCCAGAAACATGACCTGAGCGA 60
QY 61 AGCCAAGCCGGCGGCGAGGTGGCTTTGATAGTGGTGGTGCACATTCCTGGCCTTGA 120
DB 61 AGCCAAGCCGGCGGCGAGGTGGCTTTGATAGTGGTGGTGCACATTCCTGGCCTTGA 120
QY 121 TGAGCCGTA CGCCTCTGTAAACCAACTTCTCACTTTGAAACAGCTGCTGTTGAGC 180
DB 121 TGAGCCGTA CGCCTCTGTAAACCAACTTCTCACTTTGAAACAGCTGCTGTTGAGC 180
QY 181 ATTAATGAAGATTAGTCAGTGACAGGCTGGTGTGAGTCCGCCACATAGAGAATCAA 240
DB 181 ATTAATGAAGATTAGTCAGTGACAGGCTGGTGTGAGTCCGCCACATAGAGAATCAA 240
QY 241 AAATGTCCAAATGTAACTGGAGAGAAAGTGGGCACTTTTGGAGTGACATTTTCCACAGG 300
DB 241 AAATGTCCAAATGTAACTGGAGAGAAAGTGGGCACTTTTGGAGTGACATTTTCCACAGG 300
QY 301 AACTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGAGGACAGCCCCAG 360
DB 301 AACTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGAGGACAGCCCCAG 360
QY 361 TGAAGCACAGGAAGTGGACCCCCCACCCTGGGCCACCCAAACCTGGACACGTTTACCCC 420
DB 361 TGAAGCACAGGAAGTGGACCCCCCACCCTGGGCCACCCAAACCTGGACACGTTTACCCC 420
QY 421 GGAGGAGCTGCTGCAGCAGATGAAGAGCTCTCTGACCAAGAACCCACAGCTGAAGAGC 480
DB 421 GGAGGAGCTGCTGCAGCAGATGAAGAGCTCTCTGACCCGAGAACCCACAGCTGAAGAGC 480
QY 481 CATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTGAGGAGCTTTTGGCCTTGGACAGA 540
DB 481 CATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTGAGGAGCTTTTGGCCTTGGACAGA 540
QY 541 GAAACAGAGGAAGAACGCGAGTTTTTTTGAGATACAGACAAAGAACAAAGAGCGTCT 600
DB 541 GAAACAGAGGAAGAACGCGAGTTTTTTTGAGATACAGAGCAAGAACAAAGAGCGTCT 600
QY 601 AATGGCCTTGAGTCATGAGAAATGAGAAATTTGAAGGAGAGCTTGGAAATCAAAAGGAA 660
DB 601 AATGGCCTTGAGTCATGAGAAATGAGAAATTTGAAGGAGAGCTTGGAAATCAAAAGGAA 660
QY 661 ATCAGAAAGGTCTATCTGAGAACCCCACTGATGACTCCAGGCTTTCCAGGCGCGAAGCGGA 720
DB 661 ATCAGAAAGGTCTATCTGAGAACCCCACTGATGACTCCAGGCTTTCCAGGCGCGAAGCGGA 720
QY 721 GCAGGAAAGGACACAGTCCAGGACCCAGGTGTTGAGGCTACAGACAGAGACACCT 780
DB 721 GCAGGAAAGGACACAGTCCAGGACCCAGGTGTTGAGGCTACAGACAGAGACACCT 780
QY 781 GTTGGCATCGTCTGTAATCGAGCTCAAGCTGAACTCCAGCGGCTCTCAGAGATTTC 840
DB 781 GTTGGCATCGTCTGTAATCGAGCTCAAGCTGAACTCCAGCGGCTCTCAGAGATTTC 840

QY 841 CTTTGTGAAATTAGATGCTCAAGAGAGCAGAGGCTCAGTAAAGAAATCAAGCA 900
DB 841 CTTTGTGAAATTAGATGCTCAAGAGAGCAGAGGCTCAGTAAAGAAATCAAGCA 900
QY 901 TAGTCTCTGGGCCCCAGAGAAACAGTCTCCACTGGCAGCGCATTTGCTAAATATAGGAGCAG 960
DB 901 TAGTCTCTGGGCCCCAGAGAAACAGTCTCCACTGGCAGCGCATTTGCTAAATATAGGAGCAG 960
QY 961 ATCTGCAGATGGGGCCCAAGAAATTTACTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCCT 1020
DB 961 ATCTGCAGATGGGGCCCAAGAAATTTACTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCCT 1020
QY 1021 GCTGTGCTTAAGGGAAGGGAATCAGAGGTGGAGAGACTTGAAGTTGCACCTCAAGGAGGC 1080
DB 1021 GCTGTGCTTAAGGGAAGGGAATCAGAGGTGGAGAGACTTGAAGTTGCACCTCAAGGAGGC 1080
QY 1081 CAAAGAAAGATTTCAGATTTTGAAGAAAGAAAAAAGTAATCTGTTCTGAGATTGAACCCCA 1140
DB 1081 CAAAGAAAGATTTCAGATTTTGAAGAAAGAAAAAAGTAATCTGTTCTGAGATTGAACCCCA 1140
QY 1141 GACAGAGGGGAGCACAGAGAAAGAGAAATGATGAAGAGAGGCCCGGAGACTGTTGGAAG 1200
DB 1141 GACAGAGGGGAGCACAGAGAAAGAGAAATGATGAAGAGAGGCCCGGAGACTGTTGGAAG 1200
QY 1201 CGAAGTGGAGCACTGAAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCA 1260
DB 1201 CGAAGTGGAGCACTGAAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCA 1260
QY 1261 TACAAAACCTCAGGGAAGCTGAGCTAATGAAGAAAGAGACTTCAAGAAAGTGTGAGGCCCT 1320
DB 1261 TACAAAACCTCAGGGAAGCTGAGCTAATGAAGAAAGAGACTTCAAGAAAGTGTGAGGCCCT 1320
QY 1321 TGAAGGAAAAATTTCTGCAATTCATCAGAGTTTGAATGAAAGCAAGAGCTTCTTTATAC 1380
DB 1321 TGAAGGAAAAATTTCTGCAATTCATCAGAGTTTGAATGAAAGCAAGAGCTTCTTTATAC 1380
QY 1381 TAACAAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGAAACAGGC 1440
DB 1381 TAACAAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGAAACAGGC 1440
QY 1441 TAAACAGAGGATGAAGAGTCCAAATTAATCTGCTGCTACAGATGACACACAAAGCTTCT 1500
DB 1441 TAAACAGAGGATGAAGAGTCCAAATTAATCTGCTGCTACAGATGACACACAAAGCTTCT 1500
QY 1501 TCAAGAACATAATAATGCAATTCGAACCAATTCAGGAACTTAAAGAAAGAGTCAAGAAA 1560
DB 1501 TCAAGAACATAATAATGCAATTCGAACCAATTCAGGAACTTAAAGAAAGAGTCAAGAAA 1560
QY 1561 AGTGGACAGGCGAGTCTGAAGGAACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCT 1620
DB 1561 AGTGGACAGGCGAGTCTGAAGGAACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCT 1620
QY 1621 GGTTCCTCAACAGCTGCAAAATGGAATGAAGCAAAACATTTGCCAAGCAGGAAGAGGA 1680
DB 1621 GGTTCCTCAACAGCTGCAAAATGGAATGAAGCAAAACATTTGCCAAGCAGGAAGAGGA 1680
QY 1681 CTTGGAACCATGACCATCTCAGGCTCAGATGGAAGTTACTGTTCTGATTTTATGC 1740
DB 1681 CTTGGAACCATGACCATCTCAGGCTCAGATGGAAGTTACTGTTCTGATTTTATGC 1740
QY 1741 TGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCACTTGAGCTGGC 1800
DB 1741 TGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCACTTGAGCTGGC 1800
QY 1801 AGTTCTGTGAAAGAGATGATGCTTTTGAAGACCGAGGACGCGAGCTCTTGTATGAGAT 1860
DB 1801 AGTTCTGTGAAAGAGATGATGCTTTTGAAGACCGAGGACGCGAGCTCTTGTATGAGAT 1860
QY 1861 GCAGAGTCTGATGGGGGAGAACAGTGACTCTGACAGAGGCTTACTCTGTTTCAAG 1920
DB 1861 GCAGAGTCTGATGGGGGAGAACAGTGACTCTGACAGAGGCTTACTCTGTTTCAAG 1920
QY 1921 AGGAGCTGAGGACAGGAGACTGGCGCAACAGCGGAATATTCGAGTTTCACTTCTGCCCCAA 1980

```

Db      1921 AGGAGCTGAGGACAGGACTGGCGCACAGCGGAATATTCGGATTCATCTCGCCCAA 1980
      |||
Qy      1981 GGTGTGAGAGGTTCTGCCTGCATACACACAGTTACAGATTACAGTTGATGATTCATCAT 2040
      |||
Db      1981 GGTGTGAGAGGTTCTGCCTGCATACACACAGTTACAGATTACAGTTGATGATTCATCAT 2040
      |||
Qy      2041 TTAAGTGTGTGATGATACCTCCCAAACTGTGGT 2077
      |||
Db      2041 TTAAGTGTGTGATGATACCTCCCAAACTGTGGT 2077
      |||
RESULT 2
ADO43183
ID      ADO43183 standard; cDNA; 2077 BP.
XX      ADO43183;
XX
XX      29-JUL-2004 (first entry)
XX      Human optineurin isoform 1 cDNA, associated with glaucoma.
XX
XX      Human; optineurin; glaucoma; diagnosis; gene therapy; ophthalmological;
KW      gene; chromosome 10p14; ss.
XX
XX      Homo sapiens.
XX
XX      Key
XX      CDS
XX      Location/Qualifiers
XX      311..2044
XX      /tag= a
XX      /product= "Human optineurin isoform 1"
XX      replace(356,g)
XX      /tag= b
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, P16A"
XX      replace(412,a)
XX      /tag= c
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Does not alter encoded amino acid (Thr-34)"
XX      replace(432,c)
XX      /tag= d
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, Q42H"
XX      replace(433,a)
XX      /tag= e
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Does not alter encoded amino acid (Leu-41)"
XX      replace(458,a)
XX      /tag= f
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted protein change, E50K; disease-causing
XX      alteration"
XX      replace(585,t)
XX      /tag= g
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, E32V"
XX      replace(603,a)
XX      /tag= h
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted protein change, M98K; risk-associated
XX      alteration"
XX      replace(691..692,tagg)
XX      /tag= i
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Premature stop; disease-causing alteration"
XX      replace(799,g)
XX      /tag= j
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Does not alter encoded amino acid (Glu-163)"
XX      /tag= k
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, S201P"
XX
XX      variation
XX      replace(947,a)
XX      /tag= l
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, H213K"
XX      replace(949,a)
XX      /tag= m
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, H213K"
XX      replace(958,c)
XX      /tag= n
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, R216S"
XX      replace(1045,a)
XX      /tag= o
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Does not alter encoded amino acid (Gln-245)"
XX      replace(1201,t)
XX      /tag= p
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Does not alter encoded amino acid (Ser-297)"
XX      replace(1274,a)
XX      /tag= q
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, E322K"
XX      replace(1379,a)
XX      /tag= r
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted amino acid change, P357T"
XX      replace(1944,a)
XX      /tag= s
XX      /standard_name= "Single nucleotide polymorphism"
XX      /note= "Predicted protein change, R545Q; disease-causing
XX      alteration "
XX      WO2004039312-A2.
XX      13-MAY-2004.
XX
XX      26-JUN-2003; 2003WO-US020165.
XX
XX      25-OCT-2002; 2002US-00281457.
XX
XX      (UYCO-) UNIV CONNECTICUT.
XX      (SGEO-) ST GEORGES ENTERPRISES LTD.
XX
XX      Sarfarazi M, Rezaie T, Child AH;
XX
XX      WPI; 2004-376046/35.
XX      P-PSDB; ADO43184.
XX      GENBANK; AF420371.
XX
XX      New optineurin nucleic acids and polypeptides having sequence
XX      alterations, which indicates the presence of an optineurin-associated
XX      (risk of) glaucoma, useful for detecting or diagnosing glaucoma, e.g.
XX      primary open angle glaucoma.
XX
XX      Claim 1; SEQ ID NO 1; 110pp; English.
XX
XX      The present sequence is that of cDNA encoding isoform 1 of human
XX      optineurin (optic neuropathy inducing protein). Mutation of the
XX      optineurin gene on chromosome 10 is associated with primary open angle
XX      glaucoma. In specific embodiments of the invention, the alteration is: a
XX      change from GAG to AAG at codon 50 of the optineurin gene; an insertion
XX      of AG after codon 127; a change from CGG to CAG at codon 545; or a
XX      combination of one or more of these. These alterations are associated
XX      with glaucoma, and the presence of one or more of these alterations is a
XX      diagnostic for glaucoma. In another embodiment, the alteration is a
XX      change from ATG to AAG at codon 98. The presence of this alteration is
XX      indicative of an increased risk of glaucoma, and is diagnostic of this
XX      increased risk. Other alterations include a change from CCC to GCC at
XX      codon 16, a change from CAG to CAC at codon 42, a change from GAA to GTA
XX      at codon 92, and a change from GAA to AAA at codon 322. Alterations
XX      include combinations of alterations such as combinations of alterations

```

CC associated with glaucoma and those associated with an increased risk of
CC glaucoma. Methods of detection, prognosis and diagnosis of the presence
CC or absence of optineurin-associated glaucoma or of an optineurin-
CC associated increased risk of glaucoma are described, in which a sample is
CC tested for the presence of these gene sequence alterations or for
CC alterations in the expression or activity of the optineurin protein. Also
CC described are methods of therapy of glaucoma, including gene therapy
CC methods.

XX Sequence 2077 BP; 682 A; 431 C; 566 G; 398 T; 0 U; 0 Other;

Query Match	99.98;	Score 2075.4;	DB 12;	Length 2077;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 2076;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy 1	ATCCCGTGGGAGTTCTCTCCAGCGGCACGATCCGAGGAACAGTGAACCTTGAGCGA	60		
Db 1	ATCCCGTGGGAGTTCTCTCCAGCGGCACGATCCGAGGAACAGTGAACCTTGAGCGA	60		
Qy 61	AGCCAAGCCGGCGGCAAGTGTGGCTTTGATAGCTGGTGTGCGCACTTCTGGCCTTGA	120		
Db 61	AGCCAAGCCGGCGGCAAGTGTGGCTTTGATAGCTGGTGTGCGCACTTCTGGCCTTGA	120		
Qy 121	TGAGCCGTACGCCCTCTGTAAACCCAACTTCTCACTTTGAAACAGCTGGTTGAGC	180		
Db 121	TGAGCCGTACGCCCTCTGTAAACCCAACTTCTCACTTTGAAACAGCTGGTTGAGC	180		
Qy 181	ATTAAATGAAGATTAGTCAGTGACAGGCTGTGTGCTGAGTCCGCACATAGAGAATCAA	240		
Db 181	ATTAAATGAAGATTAGTCAGTGACAGGCTGTGTGCTGAGTCCGCACATAGAGAATCAA	240		
Qy 241	AAATGTCCAAAATGTAACTGAGAGAAAGTGGCAACTTTTGGAGTGACTTTTCCACAGG	300		
Db 241	AAATGTCCAAAATGTAACTGAGAGAAAGTGGCAACTTTTGGAGTGACTTTTCCACAGG	300		
Qy 301	AACCTTGTGCAATGCCATCAACCTCTCAGCTGGTCACTGAAAGAGAGACAGCCCGAG	360		
Db 301	AACCTTGTGCAATGCCATCAACCTCTCAGCTGGTCACTGAAAGAGAGACAGCCCGAG	360		
Qy 361	TGAAGCACAGGAAATGGACCCGCCACCTGGGCCACCCCAAACTGGACACGTTTACCCC	420		
Db 361	TGAAGCACAGGAAATGGACCCGCCACCTGGGCCACCCCAAACTGGACACGTTTACCCC	420		
Qy 421	GGAGGAGCTCTGCAGCAGATGAAAGAGCTCTGACCCAGAACCCACAGCTGAAAGAGC	480		
Db 421	GGAGGAGCTCTGCAGCAGATGAAAGAGCTCTGACCCAGAACCCACAGCTGAAAGAGC	480		
Qy 481	CATGAAGCTAAATCAACCCATGAAGGGAGATTGAGAGCTTTGGCCTTGACAGA	540		
Db 481	CATGAAGCTAAATCAACCCATGAAGGGAGATTGAGAGCTTTGGCCTTGACAGA	540		
Qy 541	GAACACAGAGGAAGAGCCAGTTTTTTTGAGATACAGACCAAGAGCAAGAGCGTCT	600		
Db 541	GAACACAGAGGAAGAGCCAGTTTTTTTGAGATACAGACCAAGAGCAAGAGCGTCT	600		
Qy 601	AATGGCCTTGAGTCATGAGAAATGAGAAATGAAAGGAGAGCTTTGGAAAACCTAAAGGGA	660		
Db 601	AATGGCCTTGAGTCATGAGAAATGAGAAATGAAAGGAGAGCTTTGGAAAACCTAAAGGGA	660		
Qy 661	ATCAGAAAGTCACTTGAGACCCCACTGATGATCAAGCTTCCAGGCGCCGAAAGCGGA	720		
Db 661	ATCAGAAAGTCACTTGAGACCCCACTGATGATCAAGCTTCCAGGCGCCGAAAGCGGA	720		
Qy 721	GCAGGAAGGACAGCTCAGGACCCAGTGGTGGCTACAGACAGAGAGGACAGCT	780		
Db 721	GCAGGAAGGACAGCTCAGGACCCAGTGGTGGCTACAGACAGAGAGGACAGCT	780		
Qy 781	GTTGGGCATCGTGTCTGAACTGACAGCTCAAGCTGAACTCCAGCGGCTCTCAGAGATTC	840		
Db 781	GTTGGGCATCGTGTCTGAACTGACAGCTCAAGCTGAACTCCAGCGGCTCTCAGAGATTC	840		
Qy 841	CTTTGTTGAAATAGGATGGCTGAAGAGAGAGAGAGGATGAGTAAAGAAATCAAGCA	900		
Db 841	CTTTGTTGAAATAGGATGGCTGAAGAGAGAGAGAGGATGAGTAAAGAAATCAAGCA	900		

841	CTTTGTTGAAATAGGATGGCTGAAGAGAGAGAGGATGAGTAAAGAAATCAAGCA	900
901	TAGTCTCGGCCCCACAGAGAACAGTCTCCACTGCACGGCAATGCTCTAAATATAGGACGAG	960
901	TAGTCTCGGCCCCACAGAGAACAGTCTCCACTGCACGGCAATGCTCTAAATATAGGACGAG	960
961	ATCTGAGATGGGGCCAAAGAAATTAATCTGAAACATGAGGAGTTAACTGTGAGCAGCTCCT	1020
961	ATCTGAGATGGGGCCAAAGAAATTAATCTGAAACATGAGGAGTTAACTGTGAGCAGCTCCT	1020
1021	GCTGTGCTTAAGGAGAGGAGNATCAGAGGTGGAGAGACTTTGAAGTTGCACCTCAAGGAGGC	1080
1021	GCTGTGCTTAAGGAGAGGAGNATCAGAGGTGGAGAGACTTTGAAGTTGCACCTCAAGGAGGC	1080
1081	CAAAAGAAAGAGTTTCAGATTTTCAAAAGAAAAAACAAGTAATCGTTCTGAGATTTGAAACCCA	1140
1081	CAAAAGAAAGAGTTTCAGATTTTCAAAAGAAAAAACAAGTAATCGTTCTGAGATTTGAAACCCA	1140
1141	GACAGAGGGAGCACAGAGAAAGAAATGATGAAGAGAAAAAGGCCCGGAGACTGTTGGAAG	1200
1141	GACAGAGGGAGCACAGAGAAAGAAATGATGAAGAGAAAAAGGCCCGGAGACTGTTGGAAG	1200
1201	CGAAGTGGAAAGCACTGAACTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCA	1260
1201	CGAAGTGGAAAGCACTGAACTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCA	1260
1261	TACAAAACCTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGGTGTCAGGCCCT	1320
1261	TACAAAACCTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGGTGTCAGGCCCT	1320
1321	TGAAGGAAAAATTTCTGCAATTCATCAGAGTTGAAATGAAAAGCAAGAGCTTGTTTATAC	1380
1321	TGAAGGAAAAATTTCTGCAATTCATCAGAGTTGAAATGAAAAGCAAGAGCTTGTTTATAC	1380
1381	TAACAAAAAGTTAGAGCTACAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGC	1440
1381	TAACAAAAAGTTAGAGCTACAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGC	1440
1441	TAAAAACAGAGGATGAAAAAGTCCAAAATTAACCTGCTACAGATCACAACAAAGCTTCT	1500
1441	TAAAAACAGAGGATGAAAAAGTCCAAAATTAACCTGCTACAGATCACAACAAAGCTTCT	1500
1501	TCAAGAACATAATTAATGCAATTTGAAAAACAATTGAGGAACTAAACAAGAAAAAGAGTCA	1560
1501	TCAAGAACATAATTAATGCAATTTGAAAAACAATTGAGGAACTAAACAAGAAAAAGAGTCA	1560
1561	AGTGGACAGGCGAGTCTGAAAGAACTGAGTGAAGAAAACTGGAACTGGCAGAGAGGCTCT	1620
1561	AGTGGACAGGCGAGTCTGAAAGAACTGAGTGAAGAAAACTGGAACTGGCAGAGAGGCTCT	1620
1621	GGCTTCCAAACAGCTGCAAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT	1680
1621	GGCTTCCAAACAGCTGCAAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT	1680
1681	CCTGGAACCATACCATCTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCTATGC	1740
1681	CCTGGAACCATACCATCTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCTATGC	1740
1741	TGAAGAGCAGCAGAGAGAAAAATTCATGAGGAAAAAGGAGCAACTGGCAATGAGCTGGC	1800
1741	TGAAGAGCAGCAGAGAGAAAAATTCATGAGGAAAAAGGAGCAACTGGCAATGAGCTGGC	1800
1801	AGTTCTGCTGAAGAGAAATGATGCTTTTCAAGACGAGGAGGAGCTCTTGTATGAGAT	1860
1801	AGTTCTGCTGAAGAGAAATGATGCTTTTCAAGACGAGGAGGAGCTCTTGTATGAGAT	1860
1861	GCAGAGTCTCTCATGGGGCGAGAAACAGTGAATCTGACACAGAGGCTTACCTCTTCAAAG	1920
1861	GCAGAGTCTCTCATGGGGCGAGAAACAGTGAATCTGACACAGAGGCTTACCTCTTCAAAG	1920
1921	AGGAGCTGAGGACAGGGAATGGCGGCAACAGCGGAATATTCGATTCATTTCTGCCCCAA	1980
1921	AGGAGCTGAGGACAGGGAATGGCGGCAACAGCGGAATATTCGATTCATTTCTGCCCCAA	1980

QY	840	CTTTTGTGAAATTAGGATGGCTGAAGAGGAAGCAGAGGGTCACTAGTAAAGAAATCAAGC	899	1920	GAGGAGCTGAGACAGGAGCTGGCGGCAACAGCGGAATATTCGGATTCAATTCCTGCCCA	1979
Db	856	CTTTTGTGAAATTAGGATGGCTGAAGAGGAAGCAGAGGGTCACTAGTAAAGAAATCAAGC	915	1936	GAGGAGCTGAGACAGGAGCTGGCGGCAACAGCGGAATATTCGGATTCAATTCCTGCCCA	1995
QY	900	ATAGTCTGGGCCACGAGAAAGTCTCACTGGCAGCGCATTTCTAAATATAGAGCA	959	1980	AGTGTGGAGAGGTTCTGCTGCATAGACACAGTTCACAGATTACCGTGTGGATTGCATCA	2039
Db	916	ATAGTCTGGGTCACGAGAACAGTCTCCACTGGCAGCGCATTTCTCATATAGAGGA	975	1996	AGTGTGGAGAGGTTCTGCTGCATAGACACAGTTCACAGATTACCGTGTGGATTGCATCA	2055
QY	960	GATCTCAGATGGGCCAAGAAATTAATTCGAAATCAGGAGTTAACTGTGAGCCAGCTCC	1019	2040	TTTAAAGTGTGATGATPACACCTCCCAAACTGTTGGT	2077
Db	976	GATCTCAGATGGGCCAAGAAATTAATTCGAAATCAGGAGTTAACTGTGAGCCAGCTCC	1035	2056	TTTAAAGTGTGATGATPACACCTCCCAAACTGTTGGT	2093
QY	1020	TGCTGTGCCCTAAGGGAAGGGAATCAGAAGGTGGAGAGACTTGAAGTTGCACCTCAAGGAG	1079	RESULT 6		
Db	1036	TGCTGTGCCCTAAGGGAAGGGAATCAGAAGGTGGAGAGACTTGAAGTTGCACCTCAAGGAG	1095	ADR24267		
QY	1080	CMAAGAAAGATTTCAGATTGTAAGAAAGAAACAAGTAATCGTTCTGAGATTGAACCC	1139	ID	ADR24267 standard; DNA; 2139 BP.	
Db	1096	CMAAGAAAGATTTCAGATTGTAAGAAAGAAACAAGTAATCGTTCTGAGATTGAACCC	1155	XX	ADR24267;	
QY	1140	AGACAGAGGGAGCAGAGAGAAAGAAATGATGAAGAGAAAGCCCGGAGACTGTTGGAA	1199	XX	21-OCT-2004 (first entry)	
Db	1156	AGACAGAGGGAGCAGAGAGAAAGAAATGATGAAGAGAAAGCCCGGAGACTGTTGGAA	1215	XX	Breast cancer prognosis marker #128.	
QY	1200	GCAGTGGAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTC	1259	XX	ds; breast cancer; prognosis; gene expression; diagnosis.	
Db	1216	GCAGTGGAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTC	1275	XX	Homo sapiens.	
QY	1260	ATCAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAGGCC	1319	XX	WO2004065545-A2.	
Db	1276	ATCAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAGGCC	1335	XX	05-AUG-2004.	
QY	1320	TTCAAGAGAAATCTGCAATTCATCAGATTGAATGAAGCAAGAGCTTGTATTATA	1379	XX	15-JAN-2004; 2004WO-US001100.	
Db	1336	TTCAAGAGAAATCTGCAATTCATCAGATTGAATGAAGCAAGAGCTTGTATTATA	1395	XX	15-JAN-2003; 2003US-00342887.	
QY	1380	CTAACAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGG	1439	XX	(ROSE-) ROSETTA INPHARMATICS LLC.	
Db	1396	CTAACAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGG	1455	XX	(NECA-) NETHERLANDS CANCER INST.	
QY	1440	CTAACACAGAGGATGAAGTCAAAATTAATCTGTCTACGATGACACACAAAGCTTC	1499	XX	Van't Veer LJ, He Y;	
Db	1456	CTAACACAGAGGATGAAGTCAAAATTAATCTGTCTACGATGACACACAAAGCTTC	1515	XX	WPI; 2004-593473/57.	
QY	1500	TTCAAGAACATAATATGCTTGAACCAATTCAGGAATTAACAAGAAAGAGTTCAGAAA	1559	XX	Classifying a breast cancer patient according to prognosis comprises	
Db	1516	TTCAAGAACATAATATGCTTGAACCAATTCAGGAATTAACAAGAAAGAGTTCAGAAA	1575	XX	determining the similarity between the level of expression of each of	
QY	1560	AAGTGGACAGGCGAGTCTGAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTC	1619	XX	five genes in a cell sample taken from patient, to control levels.	
Db	1576	AAGTGGACAGGCGAGTCTGAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTC	1635	XX	Disclosure; SEQ ID NO 128; 226pp; English.	
QY	1620	TGGCTTCAAACAGCTGCAAAATGGAATGAAGCAACCATTTGCCACGACGAGAGG	1679	XX	The invention relates to a method of classifying a breast cancer patient	
Db	1636	TGGCTTCAAACAGCTGCAAAATGGAATGAAGCAACCATTTGCCACGACGAGAGAGG	1695	XX	according to prognosis by determining the similarity between the level of	
QY	1680	ACCTGGAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTCTGATTTTCATG	1739	XX	expression of each of five genes for which markers are listed in the	
Db	1696	ACCTGGAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTCTGATTTTCATG	1755	XX	specification, in a cell sample taken from the breast cancer patient, to	
QY	1740	CTGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGAGCAACTGGCATTCAGCTGG	1799	XX	control levels of expression for each respective five genes to obtain a	
Db	1756	CTGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGAGCAACTGGCATTCAGCTGG	1815	XX	patient similarity value. The methods are useful for classifying a breast	
QY	1800	CAGTTCTGCTGAAGAGAAATGATGCTTTCGAAGCGAGGAGGAGTCTTCATGAGGA	1859	XX	cancer patient according to prognosis. Kits and computer program products	
Db	1816	CAGTTCTGCTGAAGAGAAATGATGCTTTCGAAGCGAGGAGGAGTCTTCATGAGGA	1875	XX	are useful for data analysis using the diagnostic, prognostic and	
QY	1860	TGCAGAGTCTGATGGGGCGAGAAACAGTGAATCTGACACAGGAGGCTTACCTTGTTC	1919	XX	statistical methods of the invention. This sequence corresponds to a	
Db	1876	TGCAGAGTCTGATGGGGCGAGAAACAGTGAATCTGACACAGGAGGCTTACCTTGTTC	1935	XX	marker used in the method of the invention.	
Query Match 99.0%; Score 2056.4; DB 13; Length 2139;						
Best Local Similarity 99.7%; Pred. No. 0;						
Matches 2071; Conservative 0; Mismatches 6; Indels 1; Gaps 1;						
QY	1	ATCCCCGTGGAGTTCTCTCCAGGCGGACAGTCCGAGGAAACAGTGACCTGAGCGA	60	QY	1	ATCCCCGTGGAGTTCTCTCCAGGCGGACAGTCCGAGGAAACAGTGACCTGAGCGA
Db	16	ATCCCCGTGGAGTTCTCTCCAGGCGGACAGTCCGAGGAAACAGTGACCTGAGCGA	75	Db	16	ATCCCCGTGGAGTTCTCTCCAGGCGGACAGTCCGAGGAAACAGTGACCTGAGCGA
QY	61	AGCCAAAGCCGGCGGAGGTGTGGCTTTGATAGTGGTGGTGGTGGTGGTGGTGGTGG	120	QY	61	AGCCAAAGCCGGCGGAGGTGTGGCTTTGATAGTGGTGGTGGTGGTGGTGGTGGTGG
Db	76	AGCCAAAGCCGGCGGAGGTGTGGCTTTGATAGTGGTGGTGGTGGTGGTGGTGGTGG	135	Db	76	AGCCAAAGCCGGCGGAGGTGTGGCTTTGATAGTGGTGGTGGTGGTGGTGGTGGTGG

QY 121 TGAGCGCTGACCTCTGTAAACCCAACTTCTCTCACTTTGAAACAGCTGCTGTGTCAGC 180
Db 136 TGAGCGGTACGCTCTGTAAACCCAACTTCTCTCACTTTGAAACAGCTGCTGTGTCAGC 195
QY 181 ATTAATGAAGATTAGTTCAGTGACAGGCTGTGTGCTGAGTCCGCACATAGAAGAAATCAA 240
Db 196 ATTAATGAAGATTAGTTCAGTGACAGGCTGTGTGCTGAGTCCGCACATAGAAGAAATCAA 255
QY 241 AAATGTCCAAATGTAACTGGAGAGAAAGTGGCAACTTTT - GGAGTGACTTTTCCACAG 299
Db 256 AAATGTCCAAATGTAACTGGAGAGAAAGTGGCAACTTTTGGGAGTGACTTTTCCACAG 315
QY 300 GAACCTTCTGCAATGCTCCATCAACTCTCTCAGCTGCTCTCACTGAAAAGAGAGACGCCCA 359
Db 316 GAACCTTCTGCAATGCTCCATCAACTCTCTCAGCTGCTCTCACTGAAAAGAGAGACGCCCA 375
QY 360 GTGAAAGCACAGGAATGGACCCGCCACCTGGGCCCAACCCCACTGGACACGTTTACCC 419
Db 376 GTGAAAGCACAGGAATGGACCCGCCACCTGGGCCCAACCCCACTGGACACGTTTACCC 435
QY 420 CGGAGGAGCTGCTCAGCAGATGAAGAGCTCTCTGACCAAGAACCCACAGCTGAAAGAAG 479
Db 436 CGGAGGAGCTGCTCAGCAGATGAAGAGCTCTCTGACCGAGAACCCACAGCTGAAAGAG 495
QY 480 CCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTGGGCTTGGACAG 539
Db 496 CCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTGGGCTTGGACAG 555
QY 540 AGAAACAGAGAGGAGAGCGGCTTTTTCAGATACAGACCAAGAGCAAGAGAGCGTC 599
Db 556 AGAAACAGAGAGGAGAGCGGCTTTTTCAGATACAGACCAAGAGCAAGAGAGCGTC 615
QY 600 TAATGGCTTTGAGTCATGAGAAATGAGAAATTTGAAGGAGAGCTTTGGAATACTTAAAGGGA 659
Db 616 TAATGGCTTTGAGTCATGAGAAATGAGAAATTTGAAGGAGAGCTTTGGAATACTTAAAGGGA 675
QY 660 AATCAGAAAGGTATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGGAAGCGG 719
Db 676 AATCAGAAAGGTATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGGAAGCGG 735
QY 720 AGCAGGAAAGGACCCAGCTCAGGACCCAGGTGGTGGAGCTTACAAGCAGAGAGGCGACAC 779
Db 736 AGCAGGAAAGGACCCAGCTCAGGACCCAGGTGGTGGAGCTTACAAGCAGAGAGGCGACAC 795
QY 780 TGTGTGGCATCGTGTCTGAACTGAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAGATT 839
Db 796 TGTGTGGCATCGTGTCTGAACTGAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAGATT 855
QY 840 CCTTTGTGAAATTAAGATGGCTGAAGGAGAGAGAGGCTCAGTAAAAAGAAATCAAGC 899
Db 856 CCTTTGTGAAATTAAGATGGCTGAAGGAGAGAGAGGCTCAGTAAAAAGAAATCAAGC 915
QY 900 ATAGTCTGGGCCACAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGAGCA 959
Db 916 ATAGTCTGGGTCCACAGAACAGTCTCCACTGGCAGCGCATTTGTCTCACTATAGGAGGA 975
QY 960 GATCTGCAGATGGGGCCAGAAATTTACTTTCGAAATATGAGGAGTTAACTGTGAGCCAGCTCC 1019
Db 976 GATCTGCAGATGGGGCCAGAAATTTACTTTCGAAATATGAGGAGTTAACTGTGAGCCAGCTCC 1035
QY 1020 TGCTGTGCTTAAGGGAAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGG 1079
Db 1036 TGCTGTGCTTAAGGGAAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGG 1095
QY 1080 CCAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAGTAATCGTTCTGAGATTGAACCC 1139
Db 1096 CCAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAGTAATCGTTCTGAGATTGAACCC 1155
QY 1140 AGACAGGGGGACACAGAGAAAGAGATCATGAGAGAAAGGCCCGGACACTTGTGGAA 1199
Db 1156 AGACAGGGGGACACAGAGAAAGAGATCATGAGAGAAAGGCCCGGACACTTGTGGAA 1215
QY 1200 GCGAAGTGGAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGCGTC 1259

Db 1216 GCGAAGTGGAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTC 1275
QY 1260 ATACAAACCTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAAGTGTGAGGCC 1319
Db 1276 ATACAAACCTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAAGTGTGAGGCC 1335
QY 1320 TTGAAAGGAAAAATTTCTGCAATTCATCAGAGTTGAATGAAAAGCAAGAGCTTCTTTTATA 1379
Db 1336 TTGAAAGGAAAAATTTCTGCAATTCATCAGAGTTGAATGAAAAGCAAGAGCTTCTTTTATC 1395
QY 1380 CTAACAAAAAGTTAGAGCTACAGTGGAAGCATGCTATCAGAAATCAAAATGGAACAGG 1439
Db 1396 CTAACAAAAAGTTAGAGCTACAGTGGAAGCATGCTATCAGAAATCAAAATGGAACAGG 1455
QY 1440 CTAACAAAGAGGATGAAAGTCCAAATTAACCTGTCTACAGATGACACACAAACAGCTTC 1499
Db 1456 CTAACAAAGAGGATGAAAGTCCAAATTAACCTGTCTACAGATGACACACAAACAGCTTC 1515
QY 1500 TTCAAGAACATAATAATGCAATTTGAAACCAATTTGAGGAACTTAAACAAAGAAAAAGTCAAGAA 1559
Db 1516 TTCAAGAACATAATAATGCAATTTGAAACCAATTTGAGGAACTTAAACAAAGAAAAAGTCAAGAA 1575
QY 1560 AAGTGGACAGGGCAGTGTCTGAAGAACTGAGTGAAAAAATCTGGAACCTGGCAGAGAGGCTC 1619
Db 1576 AAGTGGACAGGGCAGTGTCTGAAGAACTGAGTGAAAAAATCTGGAACCTGGCAGAGAGGCTC 1635
QY 1620 TGCTTTCCAAACAGCTGCTCAATGGAATGGAATGAGCAAAACCATTTGCCAAGCAGAGAGG 1679
Db 1636 TGCTTTCCAAACAGCTGCTCAATGGAATGGAATGAGCAAAACCATTTGCCAAGCAGAGAGG 1695
QY 1680 ACCTGGAACCAATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTTGATTTTCATG 1739
Db 1696 ACCTGGAACCAATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTTGATTTTCATG 1755
QY 1740 CTGAAAGAGCAGCGAGAGAGAAAAATTCATGAGGAAAAAGAGCAACTGGCATTGCAAGCTGG 1799
Db 1756 CTGAAAGAGCAGCGAGAGAGAAAAATTCATGAGGAAAAAGAGCAACTGGCATTGCAAGCTGG 1815
QY 1800 CAGTTCTGCTGAAAGAGAAATGATGCTTTCGAAAGCAGGAGCAGGAGCTCTTGATGAGGA 1859
Db 1816 CAGTTCTGCTGAAAGAGAAATGATGCTTTCGAAAGCAGGAGCAGGAGCTCTTGATGAGGA 1875
QY 1860 TGCAAGAGTCTGATGGGCGAGAACCAAGTCACTCTGACAGCAGGCTTACCTGTTTCAAA 1919
Db 1876 TGCAAGAGTCTGATGGGCGAGAACCAAGTCACTCTGACAGCAGGCTTACCTGTTTCAAA 1935
QY 1920 GAGGAGCTGAGGACAGGGAATGCGCGCAACAGCGGAATATTCGGAATTCATTCTGCCCCCA 1979
Db 1936 GAGGAGCTGAGGACAGGGAATGCGCGCAACAGCGGAATATTCGGAATTCATTCTGCCCCCA 1995
QY 1980 AGTGTGAGAGGTTCTGCTGACATAGACACAGTTACAGATTCAGTGATGATGATGATCA 2039
Db 1996 AGTGTGAGAGGTTCTGCTGACATAGACACAGTTACAGATTCAGTGATGATGATGATGATCA 2055
QY 2040 TTTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2077
Db 2056 TTTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2093

RESULT 7
ADL24738

ID ADL24738 standard; DNA; 2108 BP.

XX ADL24738;

XX 20-MAY-2004 (first entry)

XX Intestinal epithelium/peyer's patch M cell-associated DNA sequence #150.

DE Intestinal epithelium cell development; peyer's patch M cell development;
XX inflammatory bowel disease; glutenenteropathy; infectious disease;
KW autoimmune disease; haemolytic anaemia; rheumatoid arthritis; dermatitis;

KW Grave's disease; multiple sclerosis; allergy; asthma; diabetic mellitus;
KW immune system disorder; hypersensitivity; anaphylaxis;
XX blood group incompatibility; ds; human.
OS Homo sapiens.
XX WO200280852-A2.
XX 17-OCT-2002.
XX 04-APR-2002; 2002WO-US010873.
XX 04-APR-2001; 2001US-0281416P.
XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX Brayden DJ, Byrne D, O'mahony DJ, Evans CF, Mah SP, Lo DD;
XX WPI; 2003-075470/07.
XX Novel isolated or purified polypeptide encoded by genes associated with
PT intestinal epithelium or M cell development, differentiation or function,
PT useful for treating autoimmune diseases and infectious diseases.
XX Claim 1; SEQ ID NO 248; 152pp; English.
XX The invention comprises DNA sequences which are associated with
CC intestinal epithelium and Peyer's patch M cells. The DNA sequences of the
CC invention are useful for assessing, modifying, modulating or regulating
CC intestinal epithelium or M cell development. The DNA sequences of the
CC invention are also useful in the treatment of: inflammatory bowel
CC disease, glutenenteropathy, infectious diseases, autoimmune diseases
CC (e.g. haemolytic anaemia, rheumatoid arthritis, dermatitis, Grave's
CC disease, multiple sclerosis, allergy, asthma and diabetic mellitus),
CC diseases or disorders of the immune system, hypersensitivity,
CC anaphylaxis, and blood group incompatibility. The present nucleic acid
CC represents an intestinal epithelium/Peyer's patch M cell-associated DNA
CC sequence of the invention. NOTE: The present sequence is not shown in the
CC specification, but has been retrieved from the WIPO website.
XX
SQ Sequence 2108 BP; 686 A; 439 C; 573 G; 410 T; 0 U; 0 Other;

Query Match 98.9%; Score 2053.2; DB 10; Length 2108;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2069; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 ATCCGGTGGGAGTTCTCTCCAGGGGCGACGATCCGAGGAAACAGTGACCTGAGCGA 60
DB 14 ATCCGGTGGGAGTTCTCTCCAGGGGCGACGATCCGAGGAAACAGTGACCTGAGCGA 73
QY 61 AGCCAAAGCCGGGGCAG -GTGTGGCTTTGATAGCTGGTGGTCCACTTCTCGGCTTGG 119
DB 74 AGCCAAAGCCGGGGCAGAGTGTGGCTTTGATAGCTGGTGGTCCACTTCTCGGCTTGG 133
QY 120 ATGAGCGGTACGGCTGTGTAACCAACTCTCTCACTTTGAAACAGCTGCTGTTTCAG 179
DB 134 ATGAGCGGTACGGCTGTGTAACCAACTCTCTCACTTTGAAACAGCTGCTGTTTCAG 193
QY 180 CATTAATGAGATTAGTACGTGACGACGCGCTGGTGTGCTGACGACATAGAGAAATCA 239
DB 194 CATTAATGAGATTAGTACGTGACGACGCGCTGGTGTGCTGACGACATAGAGAAATCA 253
QY 240 AAAATGTCCAAAATGTAACGGAGAGAAAGTGGGCAACTTTTGGAGTGACTTTTCCACAG 299
DB 254 AAAATGTCCAAAATGTAACGGAGAGAAAGTGGGCAACTTTTGGAGTGACTTTTCCACAG 313
QY 300 GAACCTTCTGAATGTCCCATCAACCTCTAGCTGCTCTCACTGAAAGGAGGACAGCCCCA 359
DB 314 GAACCTTCTGAATGTCCCATCAACCTCTAGCTGCTCTCACTGAAAGGAGGACAGCCCCA 373
QY 360 GTGAAAGCAGAGAAATGGACCCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 419
DB 374 GTGAAAGCAGAGAAATGGACCCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 433

QY 420 CGGAGAGCTGCTGCAGCAGATGAAAGAGCTCTCTGACCAAGAACCAACAGCTGAAAGAAG 479
DB 434 CGGAGAGCTGCTGCAGCAGATGAAAGAGCTCTCTGACCAAGAACCAACAGCTGAAAGAAG 493
QY 480 CCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTTTGGAGAGCTTTTCGGCTGACAG 539
DB 494 CCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTTTGGAGAGCTTTTCGGCTGACAG 553
QY 540 AGAAAACAGAAAGAACGCCAGTTTGTGAGATACAGAGCAAAAGCAAAAGACGCTC 599
DB 554 AGAAAACAGAAAGAACGCCAGTTTGTGAGATACAGAGCAAAAGCAAAAGACGCTC 613
QY 600 TAATGGCTTGAAGTCAATGAAGATGAGAAATGAAGGAGAGCTTGAAGAACTAAAGGGA 659
DB 614 TAATGGCTTGAAGTCAATGAAGATGAGAAATGAAGGAGAGCTTGAAGAACTAAAGGGA 673
QY 660 AATCAGAAAGTCAATGAGGACCCACCTGATGACTCCAGGCTTCCAGGGCCGAAAGCGG 719
DB 674 AATCAGAAAGTCAATGAGGACCCACCTGATGACTCCAGGCTTCCAGGGCCGAAAGCGG 733
QY 720 AGCAGGAAAGAACAGCTCAGGACCCAGGTGGTGAAGCTACAAGCAGAGAGAGGACGACC 779
DB 734 AGCAGGAAAGAACAGCTCAGGACCCAGGTGGTGAAGCTACAAGCAGAGAGAGGACGACC 793
QY 780 TGTGGGCAATCGTGTCTGAAGTCAAGCTGAGCTCAAGCTCAAGCGGCTCTCAAGAGATT 839
DB 794 TGTGGGCAATCGTGTCTGAAGTCAAGCTCAAGCTCAAGCTCAAGCGGCTCTCAAGAGATT 853
QY 840 CTTTGTGTAATAGGATGGCTGAAGGAGAGAGAGAGGTCAGTAAAGAAATCAAGC 899
DB 854 CTTTGTGTAATAGGATGGCTGAAGGAGAGAGAGAGGTCAGTAAAGAAATCAAGC 913
QY 900 ATAGTCTGGGGCCACGAGAAACAGTCTCCACTGGCACCGCATTTGTCTAATATAGAGCA 959
DB 914 ATAGTCTGGGGCCACGAGAAACAGTCTCCACTGGCACCGCATTTGTCTAATATAGAGCA 973
QY 960 GATCTGAGATGGGGCAAGAAATTAATCTGAAACATGAGAGTTAACTGTGAGCCAGCTCC 1019
DB 974 GATCTGAGATGGGGCAAGAAATTAATCTGAAACATGAGAGTTAACTGTGAGCCAGCTCC 1033
QY 1020 TGTGTGCTTAAGGGAAGGGAATCAGAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGG 1079
DB 1034 TGTGTGCTTAAGGGAAGGGAATCAGAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGG 1093
QY 1080 CCAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAGACCC 1139
DB 1094 CCAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAGACCC 1153
QY 1140 AGACAGAGGGGACACAGAGAAAGAGATGATGAAGAGAAAGGCGCGGAGACTGTTGGAA 1199
DB 1154 AGACAGAGGGGACACAGAGAAAGAGATGATGAAGAGAAAGGCGCGGAGACTGTTGGAA 1213
QY 1200 GCGAGTGAAGCAGCTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTC 1259
DB 1214 GCGAGTGAAGCAGCTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTC 1273
QY 1260 ATACAAAACCTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTCAAGCCC 1319
DB 1274 ATACAAAACCTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTCAAGCCC 1333
QY 1320 TTGAAAGGAAAAATTTCTGCAATTCATCAGATTGAATGAAGAGAGAGCTTGTGTTTATA 1379
DB 1334 TTGAAAGGAAAAATTTCTGCAATTCATCAGATTGAATGAAGAGAGAGCTTGTGTTTATC 1393
QY 1380 CTAAACAAAAGTTAGAGCTACAGTGAAGAGCATGCTATCAGAAATCAAAATGGAAACAGG 1439
DB 1394 CTAAACAAAAGTTAGAGCTACAGTGAAGAGCATGCTATCAGAAATCAAAATGGAAACAGG 1453
QY 1440 CTAAAAACAGAGGATGAAAGTCCAAATTAATCTGTCTACAGATGACACACAAAGCTTC 1499
DB 1454 CTAAAAACAGAGGATGAAAGTCCAAATTAATCTGTCTACAGATGACACACAAAGCTTC 1513

QY 1500 TTCAAGAACATAATAATCATTTGAAACAAATTTGGGAACTAACAGAAAGAGTCAGAAA 1559
Db 1514 TTCAAGAACATAATAATCAATTTGAAACAAATTTGGGAACTAACAGAAAGAGTCAGAAA 1573
QY 1560 AAGTGGACAGGGCAGTCTGAGGAACTGAGTGAAGAACTGGAACCTGACAGAGAGGCTC 1619
Db 1574 AAGTGGACAGGGCAGTCTGAGGAACTGAGTGAAGAACTGGAACCTGACAGAGAGGCTC 1633
QY 1620 TGGCTTCCAAACAGCTGCAAAATGATGAATGAAGCAAAACCAATGCCCAGCAGGAAGG 1679
Db 1634 TGGCTTCCAAACAGCTGCAAAATGATGAATGAAGCAAAACCAATGCCCAGCAGGAAGG 1693
QY 1680 ACCTGGAAACCATGACCACTCTAGGGCTCAGATGGAAGTTTACGTTCGATTTTCATG 1739
Db 1694 ACCTGGAAACCATGACCACTCTAGGGCTCAGATGGAAGTTTACGTTCGATTTTCATG 1753
QY 1740 CTGAAAGCAGCAGCAGAGAGAAATTCATGAGGAAGAGGACCACTGGCATTGCGAGCTGG 1799
Db 1754 CTGAAAGCAGCAGCAGAGAGAAATTCATGAGGAAGAGGACCACTGGCATTGCGAGCTGG 1813
QY 1800 CAGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAGGAGGAGCTCTTGATGGAGA 1859
Db 1814 CAGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAGGAGGAGCTCTTGATGGAGA 1873
QY 1860 TGCAGAGTCGTATGGGCGAGAAACAAGTGACTCTGACCAAGAGGCTTACCTGTTCAAA 1919
Db 1874 TGCAGAGTCGTATGGGCGAGAAACAAGTGACTCTGACCAAGAGGCTTACCTGTTCAAA 1933
QY 1920 GAGGAGCTGAGCAGGAGGAGCTGGCGGCAACAGCGGAAATATTCGATTCATCTGCCCCA 1979
Db 1934 GAGGAGCTGAGCAGGAGGAGCTGGCGGCAACAGCGGAAATATTCGATTCATCTGCCCCA 1993
QY 1980 AGTGTGGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCACGCTGATGATGATCA 2039
Db 1994 AGTGTGGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCACGCTGATGATGATCA 2053
QY 2040 TTTAAGTGTGATGATATCACCTCCCAAACTGTTGGT 2077
Db 2054 TTTAAGTGTGATGATATCACCTCCCAAACTGTTGGT 2091

RESULT 8

ADP23349
ID ADP23349 standard; cDNA; 3904 BP.

AC ADP23349;

DT 18-NOV-2004 (first entry)

XX PRO polypeptide encoding cDNA SEQ ID NO:443.

XX ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipruritic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.

XX Unidentified.

XX WO2004041170-A2.

XX 21-MAY-2004.

XX 30-OCT-2003; 2003WO-US034312.

XX 01-NOV-2002; 2002US-0423394P.

XX (GETH) GENENTECH INC.

XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;

XX WPI; 2004-419628/39.

DR P-PSDB; ADP23350.

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

QY

Sequence 3904 BP; 1154 A; 674 C; 821 G; 928 T; 0 U; 327 Other;

Query Match

Best Local Similarity 98.8%; Score 2052.4; DB 13; Length 3904;

Matches 2075; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY	480	CCATGAAGCTAAATAATCAAGCCATCAAAAGGGAGATTTGAGGAGCTTTTCGGCTCGACAG	539	1560	AAGTCGACAGGGCAGTCTGTAAGGAACCTGAGTGAAAACTGGAACTGGCAGAGAGGCTC	1619
Db	495	CCATGAAGCTAAATAATCAAGCCATCAAAAGGGAGATTTGAGGAGCTTTTCGGCTCGACAG	554	1575	AAGTCGACAGGGCAGTCTGTAAGGAACCTGAGTGAAAACTGGAACTGGCAGAGAGGCTC	1634
QY	540	AGAAACAGAGGAGAACCCAGTCTTTTTCAGATACAGAGCAGAAAGCAAGAGCGCTC	599	1620	TGGCTTCCAAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGG	1679
Db	555	AGAAACAGAGGAGAACCCAGTCTTTTTCAGATACAGAGCAGAAAGCAAGAGCGCTC	614	1635	TGGCTTCCAAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGG	1694
QY	600	TAATGSCCTTGAGTCATGAGAAATGAGAAAATTGAAGGAAGAGCTTGGAAAACTAAAAAGGGA	659	1680	ACCTGGAAAAACCATGATACCATCTCCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTCATG	1739
Db	615	TAATGSCCTTGAGTCATGAGAAATGAGAAAATTGAAGGAAGAGCTTGGAAAACTAAAAAGGGA	674	1695	ACCTGGAAAAACCATGATACCATCTCCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTCATG	1754
QY	660	AATCAGAAAGGTCAATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCGGAAGCGG	719	1740	CTGAAAGACACGAGAGAGAAAAATTCATGAGAAAAGGAGCAACTGGCATTTGCAGCTGG	1799
Db	675	AATCAGAAAGGTCAATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCGGAAGCGG	734	1755	CTGAAAGACACGAGAGAGAAAAATTCATGAGAAAAGGAGCAACTGGCATTTGCAGCTGG	1814
QY	720	AGCAGAAAGGACACGAGCTCAGAACCCAGGTGGTGAGGCTTACAGCAGAGAGGCGAGACC	779	1800	CAGTTCTGCTGTAAGAGAAATGATGCTTTCGAAAGACGAGGAGGAGCACTCTTGATGGAGA	1859
Db	735	AGCAGAAAGGACACGAGCTCAGAACCCAGGTGGTGAGGCTTACAGCAGAGAGGCGAGACC	794	1815	CAGTTCTGCTGTAAGAGAAATGATGCTTTCGAAAGACGAGGAGGAGCACTCTTGATGGAGA	1874
QY	780	TGTTGGGCATCGTGTCTGAACTCAGCTCAAGCTCAAGCTCCAGCGGCTCCTCAGAAAGATT	839	1860	TCGAGAGTCGTCAATGGGGCGAGAAACAAAGTGACTCTGACAGCAGAGGCTTACCTTGTTCAAA	1919
Db	795	TGTTGGGCATCGTGTCTGAACTCAGCTCAAGCTCAAGCTCCAGCGGCTCCTCAGAAAGATT	854	1875	TGCAGAGTCGTCAATGGGGCGAGAAACAAAGTGACTCTGACAGCAGAGGCTTACCTTGTTCAAA	1934
QY	840	CCTTTGTTGAAATTAGATGGCTGAAAGGAGAAACAGAGGTCAGTAAAGAAATCAAGC	899	1920	GAGGAGCTGAGCAGCAGGGACTGGCGCAACAGCGGAATATTCGATTTCTCTGCCCCCA	1979
Db	855	CCTTTGTTGAAATTAGATGGCTGAAAGGAGAAACAGAGGTCAGTAAAGAAATCAAGC	914	1935	GAGGAGCTGAGCAGCAGGGACTGGCGCAACAGCGGAATATTCGATTTCTCTGCCCCCA	1994
QY	900	ATAGTCTGGGCCACGAGNACAGTCTCCACTGGCAGCGCATTTGCTTAATATAGAGCA	959	1980	AGTGTGGAGAGGTTCTGCCCTGACATAGACACGTTACAGATTACGTTGATGGATTGCATCA	2039
Db	915	ATAGTCTGGGCCACGAGNACAGTCTCCACTGGCAGCGCATTTGCTTAATATAGAGCA	974	1995	AGTGTGGAGAGGTTCTGCCCTGACATAGACACGTTACAGATTACGTTGATGGATTGCATCA	2054
QY	960	GATCTGCAGATGGGCCCAAGAAATTACTTCGAAACATGAGGAGTTAACTGTGAGCCAGCTCC	1019	2040	TTTAAAGTGTGTATGATATACCTCCCAAAAACTGTGGT	2077
Db	975	GATCTGCAGATGGGCCCAAGAAATTACTTCGAAACATGAGGAGTTAACTGTGAGCCAGCTCC	1034	2055	TTTAAAGTGTGTATGATATACCTCCCAAAAACTGTGGT	2092
QY	1020	TGCTGTGCCCTAAGGGAAGGGAATCAGAAGGTGAGAGACTTGAAGTTGCACTCAAGGAGG	1079	RESULT 9		
Db	1035	TGCTGTGCCCTAAGGGAAGGGAATCAGAAGGTGAGAGACTTGAAGTTGCACTCAAGGAGG	1094	AAH57391		
QY	1080	CCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCTGTTCTGAGATTGAACCC	1139	ID	AAH57391 standard; cdna; 3454 BP.	
Db	1095	CCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCTGTTCTGAGATTGAACCC	1154	XX	AAH57391;	
QY	1140	AGACAGAGGGGAGCACAGAGAAAGAGAAATGATGAAGAAAGGCCCGGAGACTGTTTGGAA	1199	XX	10-SEP-2001 (first entry)	
Db	1155	AGACAGAGGGGAGCACAGAGAAAGAGAAATGATGAAGAAAGGCCCGGAGACTGTTTGGAA	1214	XX	Human skeletal muscle cell specific cDNA sequence SEQ ID NO:231.	
QY	1200	GCGAAGTGGAGACCTGAACCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTC	1259	XX	Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;	
Db	1215	GCGAAGTGGAGACCTGAACCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTC	1274	XX	liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;	
QY	1260	ATACAAAACCTCAGGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAGGCC	1319	XX	metabolic disease; developmental disease; cystostatic; immunomodulatory;	
Db	1275	ATACAAAACCTCAGGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAGGCC	1334	XX	neuroprotective; gene therapy; cancer; immunopathology; neuropathology.	
QY	1320	TTCAAAAGGAAAAATTTCTGCAATTCATCAGAGTTGAATGAAAAAGCAAGACTTGTTTATA	1379	OS	Homo sapiens.	
Db	1335	TTGNAAGGAAAAATTTCTGCAATTCATCAGAGTTGAATGAAAAAGCAAGACTTGTTTATA	1394	XX	WO200132927-A2.	
QY	1380	CTAACAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAACAGG	1439	XX	10-MAY-2001.	
Db	1395	CTAACAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAACAGG	1454	XX	02-NOV-2000; 2000WO-US030396.	
QY	1440	CTAAACAGAGGATGAAAGTCCAAATTAATGCTTACAGATGACACAAACAGCTTC	1499	XX	04-NOV-1999; 99US-0163508P.	
Db	1455	CTAAACAGAGGATGAAAGTCCAAATTAATGCTTACAGATGACACAAACAGCTTC	1514	XX	(INCY-) INCYTE GENOMICS INC.	
QY	1500	TTCAAGAACATAATAATGCAATTGAAAAACAAATTGAGGAACTTAAACAGAAAAAGAGTCAGAAA	1559	XX	Sornasse T, Seilhamer JJ, Watson GA;	
Db	1515	TTCAAGAACATAATAATGCAATTGAAAAACAAATTGAGGAACTTAAACAGAAAAAGAGTCAGAAA	1574	XX	WPI; 2001-291057/30.	
				XX	New cell and tissue specific polynucleotides useful for diagnosis, prognosis or monitoring of treatments for disorders where the gene is associated with a cancer, immunopathology or neuropathology.	
				PS	Claim 1; Page 159-160; 327pp; English.	


```
xx AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC minetics, peptides, proteins, agonists, antagonists, antibodies or their
CC fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
CC agents. Expression of (I) in a sample indicates the differentiation of
CC embryonic stem cells into a tissue selected from brain, heart, kidney,
CC liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
CC to produce an expression profile that defines a metabolic or
CC developmental process, treatment, condition, disease or disorder. The
CC gene profile can be used for diagnosis, prognosis or monitoring of
CC treatments and for investigating a predisposition to a disorder where the
CC gene is associated with a cancer, immunopathology or neuropathology
xx
SQ Sequence 3454 BP; 1094 A; 725 C; 794 G; 841 T; 0 U; 0 Other;

Query Match 95.2%; Score 1976.8; DB 4; Length 3454;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 ATCCCGGTCCGGAGTTCTCCAGCGCGCACGATGCCGAGGAACAGTGAACCTTGAGCGA 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 16 ATCCCGGTCCGGAGTTCTCCAGCGCGCACGATGCCGAGGAACAGTGAACCTTGAGCGA 75
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 61 AGCCAAAGCCGGCGGAGTGTGGCTTTGATAGCTGGTGGTCCACTTTCCTGGCCTTGGGA 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 76 AGCCAAAGCCGGCGGAGTGTGGCTTTGATAGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 135
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 121 TGAGCGGTACGCTCTGTAACCAACCTTCTCACTTTGAAACAGCTGCTGCTGCTCAGC 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 136 TGAGCGGTACGCTCTGTAACCAACCTTCTCACTTTGAAACAGCTGCTGCTGCTCAGC 195
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 181 ATTAATGAGATTAGTCACTGACAGGCTGTGTGCTGAGTCCCGACATAGAAAGATCAA 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 196 ATTAATGAGATTAGTCACTGACAGGCTGTGTGCTGAGTCCCGACATAGAAAGATCAA 255
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 241 AAATGTCCAAATGTAACTGGAGAGAAAGTGGGCAACTTTTGGAGTGACTTTTCCACAG 299
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 256 AAATGTCCAAATGTAACTGGAGAGAAAGTGGGCAACTTTTGGAGTGACTTTTCCACAG 315
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 300 GAACTTCTGCAATTCGCATCAACCTCTCAGCTGCTCCTCACTGAAAGAGGACAGCCCA 359
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 316 GAACTTCTGCAATTCGCATCAACCTCTCAGCTGCTCCTCACTGAAAGAGGACAGCCCA 375
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 360 GTGAAAGCACAGGAAATGGACCCCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 418
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 376 GTGAAAGCACAGGAAATGGACCCCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 435
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 419 CCGGAGGAGCTGCTGACGAGATGAAGAGCTCTGACCCAGAGACCCAGCTGAAAGAA 478
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 436 CCGGAGGAGCTGCTGACGAGATGAAGAGCTCTGACCCAGAGACCCAGCTGAAAGAA 495
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 479 GCCATGAGCTTAAATAATCAAGCCATGAAGGAGATTTGAGGAGCTTTGGCCCTGGACA 538
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 496 GCCATGAGCTTAAATAATCAAGCCATGAAGGAGATTTGAGGAGCTTTGGCCCTGGACA 555
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 539 GAGAAACAGAGGAAGAACCGCAGTTTTTTTGGATACAGAGCAAGAAAGCAAGAGCGT 598
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 556 GAGAAACAGAGGAAGAACCGCAGTTTTTTTGGATACAGAGCAAGAAAGCAAGAGCGT 615
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 599 CTAATGCCCTTGTGCTATGAGATGAGAAATTTGAGGAAGAGCTTTGGAAGCAATTAAGGG 658
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 616 CTAATGCCCTTGTGCTATGAGATGAGAAATTTGAGGAAGAGCTTTGGAAGCAATTAAGGG 675
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 659 AAATCAGAAAGGTCACTCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGCGCGGAGCG 718
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 676 AAATCAGAAAGGTCACTCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGCGCGGAGCG 735
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 719 GAGCAGAAAGGACAGCTCAGAACCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 778
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db
Qy 736 GAGCAGGAAAGGACACAGCTCAGGACCCAGGTCGTGGGTGAGGCTTACAAGCAGAGAGGCGAC 795
Qy 779 CTGTTGGGCATCTGTCGTCGAACTCAGCTCAAGCTGAACCTCCAGCGGCTCTCTCAGAGAT 838
Db 796 CTGTTGGGCATCTGTCGTCGAACTCAGCTCAAGCTGAACCTCCAGCGGCTCTCTCAGAGAT 855
Qy 839 TCCTTTGTTGAAATTAGGATGGCTGGAAGGAGAACAGAGGGTTCAGTAAAAAGAAATCAAG 898
Db 856 TCCTTTGTTGAAATTAGGATGGCTGGAAGGAGAACAGAGGGTTCAGTAAAAAGAAATCAAG 915
Qy 899 CATAGTCTCTGGGCCCCCAGAGAACAGTCTCCATCGGCACGGCATTTCTAAATATAGGAGC 958
Db 916 CATAGTCTCTGGGCCCCCAGAGAACAGTCTCCATCGGCACGGCATTTCTAAATATAGGAGC 975
Qy 959 AGATCTCGAGATGGGGCCCAAGAAATTAATCTCGAACATCAGGAGTTAACTGTGAGCAGCCTC 1018
Db 976 AGATCTCGAGATGGGGCCCAAGAAATTAATCTCGAACATCAGGAGTTAACTGTGAGCAGCCTC 1035
Qy 1019 CTGCTGTGCTTAAGGGAAGGGAATTCAGAGGTGGAGAGCTTGAAGTTGCACTCAAGAGG 1078
Db 1036 CTGCTGTGCTTAAGGGAAGGGAATTCAGAGGTGGAGAGCTTGAAGTTGCACTCAAGAGG 1095
Qy 1079 GCCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAACTCTTCTGAGATTGAAC 1138
Db 1096 GCCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAACTCTTCTGAGATTGAAC 1155
Qy 1139 CAGACAGAGGGAGCACAGAGAAAGAAATGATGAAGAGAAAGCCCGGAGACTGTTTGA 1198
Db 1156 CAGACAGAGGGAGCACAGAGAAAGAAATGATGAAGAGAAAGCCCGGAGACTGTTTGA 1215
Qy 1199 AGCGAAGTGAAGACCTCAAGTCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCT 1258
Db 1216 AGCGAAGTGAAGACCTCAAGTCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCT 1275
Qy 1259 CATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAGTGTGAGGCC 1318
Db 1276 CATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAGAGAGACTTCAAGAAAAGTGTGAGGCC 1335
Qy 1319 CTTGAAAGGAAAAAATCTGCAATTCATCAGAGTTGAATGAATGAAGAGAGAGCTGTTGTTAT 1378
Db 1336 CTTGAAAGGAAAAAATCTGCAATTCATCAGAGTTGAATGAATGAAGAGAGAGCTGTTGTTAT 1395
Qy 1379 ACTAACAAAAGTTAGAGCTTACAAGTGAAGAGCATGCTATCAGAAAATCAAAAATGAAGACAG 1438
Db 1396 ACTAACAAAAGTTAGAGCTTACAAGTGAAGAGCATGCTATCAGAAAATCAAAAATGAAGACAG 1455
Qy 1439 GCTAAGAACAGAGATGAAGAGTCCAAATTAATCTGTGCTACAGATGACACACAAAGAGCTT 1498
Db 1456 GCTAAGAACAGAGATGAAGAGTCCAAATTAATCTGTGCTACAGATGACACACAAAGAGCTT 1515
Qy 1499 CTTCAAGAACATTAATGCAATTTGAACCAATTTGAGGAACTTAAAGAGAGAGAGTCAAGAA 1558
Db 1516 CTTCAAGAACATTAATGCAATTTGAACCAATTTGAGGAACTTAAAGAGAGAGAGTCAAGAA 1575
Qy 1559 AAAGTGAACAGGCGCAGTCTGAAGGAACTCAGTGAAGAACTTGAAGTGAAGAGAGAGCTT 1618
Db 1576 AAAGTGAACAGGCGCAGTCTGAAGGAACTCAGTGAAGAACTTGAAGTGAAGAGAGAGCTT 1635
Qy 1619 CTGGCTTCCAAA CAGCTGCAAAATGGATGAATGAAGCAAA CCAATTCGCAAGCAGGAAAGAG 1678
Db 1636 CTGGCTTCCAAA CAGCTGCAAAATGGATGAATGAAGCAAA CCAATTCGCAAGCAGGAAAGAG 1695
Qy 1679 GACCTGGAAACCATGACCATCTCAGGCTCAGATGGAGTTTACTGTTCTGATTTTCAT 1738
Db 1696 GACCTGGAAACCATGACCATCTCAGGCTCAGATGGAGTTTACTGTTCTGATTTTCAT 1755
Qy 1739 GCTGAAAGAGCAGCAGAGAGAGAAATTTATGAGGAAAAAGGAGCAA CTGGCATTTGCAGCTG 1798
Db 1756 GCTGAAAGAGCAGCAGAGAGAGAAATTTATGAGGAAAAAGGAGCAA CTGGCATTTGCAGCTG 1815
Qy 1799 GCAGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGAGCGGAGGAGCAGCTCCTTGTATGGAG 1858
Db 1816 GCAGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGAGCGGAGGAGCAGCTCCTTGTATGGAG 1875
```



```
Db      892 ATCTCCAGATGGGGCCAAAGAATTACTTGAACATAGGAGTTAACTGTGAGCCAGCTCCT 951
Qy      1021 GCTGTGCCCTAAGGGGAAGGAATCAGAAGTGTGAGAGACTTGAAGTTGCACCTCAAGGAGGC 1080
Db      952 GCTGTGCCCTAAGGGGAAGGAATCAGAAGTGTGAGAGACTTGAAGTTGCACCTCAAGGAGGC 1011
Qy      1081 CAAGAAGAGATTTCAGATTTTGAAAAGAAAACAGTAATCGTTCTTGAGATTGAACCCA 1140
Db      1012 CAAGAAGAGATTTCAGATTTTGAAAAGAAAACAGTAATCGTTCTTGAGATTGAACCCA 1071
Qy      1141 GACAGAGGGGACACAGAGAAAGAGAAATGATGAACAGAGAAAGGCCCGGAGACTTGTGGAG 1200
Db      1072 GACAGAGGGGACACAGAGAAAGAGAAATGATGAACAGAGAAAGGCCCGGAGACTTGTGGAG 1131
Qy      1201 CGAGTGAAGACATGAACCTCCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGGCTCA 1260
Db      1132 CGAAGTGAAGACATGAACCTCCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGGCTCA 1191
Qy      1261 TACAAAACCTCAGCGAAGCTGAGCTAAATCAAGAAGAGACTTCAAGAAGTGTCAAGGCCT 1320
Db      1192 TACAAAACCTCAGCGAAGCTGAGCTAAATGAAGAAGAGACTTCAAGAAGTGTCAAGGCCT 1251
Qy      1321 TGAAGGAAAAAATTCTGCAATTCCATCAGAGTTGAATGAAAAGCAAGAGCTTGTTTATAC 1380
Db      1252 TGAAGGAAAAAATTCTGCAATTCCATCAGAGTTGAATGAAAAGCAAGAGCTTGTTTATAC 1311
Qy      1381 TAACAAAAAGTTAGAGCTACAAGTGGAAAGCATCTATCAGAAATCAAAATGGAACAGGC 1440
Db      1312 TAACAAAAAGTTAGAGCTACAAGTGGAAAGCATCTATCAGAAATCAAAATGGAACAGGC 1371
Qy      1441 TAAACACAGAGATGAAAAGTCCAAATTAACTGTGTACAGATGACACACAAAGCTTCT 1500
Db      1372 TAAACACAGAGATGAAAAGTCCAAATTAACTGTGTACAGATGACACACAAAGCTTCT 1431
Qy      1501 TCAAGAACATAAATTAATGATTGAAAAACAATTGAGGAACTAACAGAAAAGAGTCAGAAAA 1560
Db      1432 TCAGAACATAAATTAATGATTGAAAAACAATTGAGGAACTAACAGAAAAGAGTCAGAAAA 1491
Qy      1561 AGTGGACAGGCGAGTGTGAAGAACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCT 1620
Db      1492 AGTGGACAGGCGAGTGTGAAGAACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCT 1551
Qy      1621 GGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAAAACCTTCCCAAGCGAGGAAGGA 1680
Db      1552 GGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAAAACCTTCCCAAGCGAGGAAGGA 1611
Qy      1681 CCTGGAACCATGATACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGC 1740
Db      1612 CCTGGAACCATGATACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGC 1671
Qy      1741 TGAAGAGCAGCGAGAGAGAAAATTTCATGAGGAAAAGGAGCAACTGGCATTGCGAGCTGGC 1800
Db      1672 TGAAGAGCAGCGAGAGAGAAAATTTCATGAGGAAAAGGAGCAACTGGCATTGCGAGCTGGC 1731
Qy      1801 AGTTCTCTGAAGAGAAATGATGCTTTGGAAGCGGAGGAGGAGCTCTTGATGGAGAT 1860
Db      1732 AGTTCTCTGAAGAGAAATGATGCTTTGGAAGCGGAGGAGGAGCTCTTGATGGAGAT 1791
Qy      1861 GCAGAGTCGTATGGGCGGAGAACAGTGAAGTCTGACCGAGGCTTACCTGTTTCAAAG 1920
Db      1792 GCAGAGTCGTATGGGCGGAGAACAGTGAAGTCTGACCGAGGCTTACCTGTTTCAAAG 1851
Qy      1921 AGGAGCTGAGGACAGGAGCTGGGCGCAACAGCGGAATATTCGGAATTCATCTCGCCCAA 1980
Db      1852 AGGAGCTGAGGACAGGAGCTGGGCGCAACAGCGGAATATTCGGAATTCATCTCGCCCAA 1911
Qy      1981 GTGTGGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCACTGATGATTGCATCAT 2040
Db      1912 GTGTGGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCACTGATGATTGCATCAT 1971
Qy      2041 TTAAGTGTGATGATATCACTCCCAAACTGTGGT 2077
Db      1972 TTAAGTGTGATGATATCACTCCCAAACTGTGGT 2008
```

RESULT 11

ADO43187
ID ADO43187 standard; cDNA; 2008 BP.

XX ADO43187;

XX 29-JUL-2004 (first entry)

XX Human optineurin isoform 3 cDNA, associated with glaucoma.

XX Human; optineurin; glaucoma; diagnosis; gene therapy; ophthalmological;
gene; chromosome 10p14; ss.

XX Homo sapiens.

FX Key Location/Qualifiers
CDS 242..1975

FT /*tag= a
FT /product= "Human optineurin isoform 3"

XX WO2004039312-A2.

XX 13-MAY-2004.

XX 26-JUN-2003; 2003WO-US020165.

XX 25-OCT-2002; 2002US-00281457.

XX (UYCO-) UNIV CONNECTICUT.

PA (SGEO-) ST GEORGES ENTERPRISES LTD.

XX Sarfarazi M, Rezaie T, Child AH;

XX WPI; 2004-376046/35.

DR P-PSDB; ADO43188.

DR GENBANK; AF420373.

XX New optineurin nucleic acids and polypeptides having sequence

alterations, which indicates the presence of an optineurin-associated
(risk of) glaucoma, useful for detecting or diagnosing glaucoma, e.g.
primary open angle glaucoma.

XX Claim 1; SEQ ID NO 5; 110pp; English.

XX The present sequence is that of cDNA encoding isoform 3 of human
optineurin (optic neuropathy inducing protein). Mutation of the
optineurin gene on chromosome 10 is associated with primary open angle
glaucoma. In specific embodiments of the invention, the alteration is
numbering according to optineurin isoform 1 cDNA ADO43183: a change
from GAG to AAG at codon 50 of the optineurin gene; an insertion of AG
after codon 127; a change from CGG to CAG at codon 545; or a combination
of one or more of these. These alterations are associated with glaucoma,
and the presence of one or more of these alterations is diagnostic for
glaucoma. In another embodiment, the alteration is a change from ATG to
AAG at codon 98. The presence of this alteration is indicative of an
increased risk of glaucoma, and is diagnostic of this increased risk.
Other alterations include a change from CCC to GCC at codon 16, a change
from CAG to CAC at codon 42, a change from GAA to GTA at codon 92, and a
change from GAA to AAA at codon 322. Alterations include combinations of
alterations such as combinations of alterations associated with glaucoma
and those associated with an increased risk of glaucoma. Methods of
detection, prognosis and diagnosis of the presence or absence of
optineurin-associated glaucoma or of an optineurin-associated increased
risk of glaucoma are described, in which a sample is tested for the
presence of these gene sequence alterations or for alterations in the
expression or activity of the optineurin protein. Also described are
methods of therapy of glaucoma, including gene therapy methods.

XX Sequence 2008 BP; 656 A; 421 C; 549 G; 382 T; 0 U; 0 Other;

Query Match

92.8%; Score 1927.4; DB 12; Length 2008;

Best Local Similarity 96.6%; Pred. No. 0; Matches 2007; Conservative 0; Mismatches 1; Indels 69; Gaps 1;			
QY	1	ATCCCCGTTGGGAGTTCTCTCCAGGCGGACACGATGCCGAGGAAACAGTGCACCTGAGCGA	60
Db	1	ATCCCCGTTGGGAGTTCTCTCCAGGCGGACACGATGCCGAGGAAACAGTGCACCTGAGCGA	60
QY	61	AGCCAAAGCGGGCGGCGAGGTGGCTTTGATAGCTGGTGGTGCACCTCTGGCCCTTGA	120
Db	61	AGCCAAAGCGGGCGGCGAGGTGGCTTTGATAGCTGGTGGTGCACCTCTGGCCCTTGA	120
QY	121	TGAGCGTACGCTCTGTGTAACCCAACTTCTCACCCTTTGAAACAGCTGGCTGGTTACG	180
Db	121	TGAGCGTACGCTCTGTGTAACCCAACTTCTCACCCTTTGAAACAGCTGGCTGGTTACG	180
QY	181	ATTAATGAGATAGTCACTGACAGGCGCTGGTGTGCTGAGTCCGCAATAGAGAATCAA	240
Db	181	ATTAATGAGATAGTCACTGACAGGCGCTGGTGTGCTGAGTCCGCAATAGAGAATCAA	240
QY	241	AAATGTCCAAAATGTAACTGGAGAGAAAGTGGGCAACTTTTGGAGTGACCTTTTCCACAGG	300
Db	231	-----G 231	
QY	301	AACCTTCGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGAGAGACAGCCCCAG	360
Db	232	AACCTTCGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGAGAGACAGCCCCAG	291
QY	361	TGAAGCACAGGAAATGGACCCCGCCACCTGGCCCGCCACCAACCTGACACGTTTACCCC	420
Db	292	TGAAGCACAGGAAATGGACCCCGCCACCTGGCCCGCCACCAACCTGACACGTTTACCCC	351
QY	421	GGAGGAGCTCTGCAGCAGATGAAGAGCTCTCGAACCAAGAACCCACAGCTGAAAGAGC	480
Db	352	GGAGGAGCTCTGCAGCAGATGAAGAGCTCTCGAACCAAGAACCCACAGCTGAAAGAGC	411
QY	481	CATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTGAGGAGCTTTGGSCCTGGACAGA	540
Db	412	CATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTGAGGAGCTTTGGSCCTGGACAGA	471
QY	541	GAACACAGAGAGAGACCCAGTTTTTTTGAGATACAGAGCAAGAGACCAAGAGCGTCT	600
Db	472	GAACACAGAGAGAGACCCAGTTTTTTTGAGATACAGAGCAAGAGACCAAGAGCGTCT	531
QY	601	AATGGCCTTCAGTCATGAGAAATGAGAAATGAAAGGAGAGCTTTGGAAAACTAAAGGGAA	660
Db	532	AATGGCCTTCAGTCATGAGAAATGAGAAATGAAAGGAGAGCTTTGGAAAACTAAAGGGAA	591
QY	661	ATCAGAAAGTCACTTGAAGACCCCACTGATGACTCAGGCTTCCGAGGCGCAAGCGGA	720
Db	592	ATCAGAAAGTCACTTGAAGACCCCACTGATGACTCAGGCTTCCGAGGCGCAAGCGGA	651
QY	721	GCAGGAAAGGACCGAGCTCAGGACCCAGGTGGTGGGCTACAGGCTACAGCAGAGAGCGACCT	780
Db	652	GCAGGAAAGGACCGAGCTCAGGACCCAGGTGGTGGGCTACAGGCTACAGCAGAGAGCGACCT	711
QY	781	GTTGGGCATCGTGTCTGAACTGACGCTCAAGCTGAACTCCAGGGCTCTCTCAGAAGATTC	840
Db	712	GTTGGGCATCGTGTCTGAACTGACGCTCAAGCTGAACTCCAGGGCTCTCTCAGAGATTC	771
QY	841	CTTTGTTGAAATTAGGATGGCTGAAGAGAGAGAGAGAGGTTCAGTAAAAAATCAAGCA	900
Db	772	CTTTGTTGAAATTAGGATGGCTGAAGAGAGAGAGAGAGGTTCAGTAAAAAATCAAGCA	831
QY	901	TAGTCTGGGCCACAGAGACGTCTCCACTGGCAGCGGCTGCTCTAAATATAGGAGCAG	960
Db	832	TAGTCTGGGCCACAGAGACGTCTCCACTGGCAGCGGCTGCTCTAAATATAGGAGCAG	891
QY	961	ATCTGCAGATGGGGCCCAAGATTAATCTCGAACTAGAGAGTTAACTGTGAGCCAGCTCTCT	1020
Db	892	ATCTGCAGATGGGGCCCAAGATTAATCTCGAACTAGAGAGTTAACTGTGAGCCAGCTCTCT	951
QY	1021	GCTGTGCTTAAGGGAAGGGAATCAGAGGTGGAGAGACTTGAAGTGTGCACCTCAAGGAGGC	1080

RESULT 12
ACF03997

Db	952	GCTGTGCTTAAGGGAAGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACCTCAAGAGGC	1011
QY	1081	CAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAATCTTCTGAGATTGAACCCCA	1140
Db	1012	CAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAATCTTCTGAGATTGAACCCCA	1071
QY	1141	GCACAGGGGAGCACAGAGAAAGAGATGATCAAGAGAAAGCCCGGAGACTGTTGGAAG	1200
Db	1072	GCACAGGGGAGCACAGAGAAAGAGATGATCAAGAGAAAGCCCGGAGACTGTTGGAAG	1131
QY	1201	CGAAGTGAAGCACCTCAAGTGCATCTCTGTTTAAAGGAGCTTCAAGAGAGCTCA	1260
Db	1132	CGAAGTGAAGCACCTCAAGTGCATCTCTGTTTAAAGGAGCTTCAAGAGAGCTCA	1191
QY	1261	TACAAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTGAGCCCT	1320
Db	1192	TACAAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTGAGCCCT	1251
QY	1321	TGAAAGGAAAAATTTGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTATTATAC	1380
Db	1252	TGAAAGGAAAAATTTGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTATTATAC	1311
QY	1381	TACAAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGC	1440
Db	1312	TACAAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGC	1371
QY	1441	TAAAAACAGAGGATGAAAAAGTCCAAATTAATCTGTGTACAGATGACACACAAAGCTTCT	1500
Db	1372	TAAAAACAGAGGATGAAAAAGTCCAAATTAATCTGTGTACAGATGACACACAAAGCTTCT	1431
QY	1501	TCAGAAACATAATATATGCAATTGAAAAATTAAGAGAACTAACAGAAAAAGTCAAGAAA	1560
Db	1432	TCAGAAACATAATATATGCAATTGAAAAATTAAGAGAACTAACAGAAAAAGTCAAGAAA	1491
QY	1561	AGTGGACAGGCGAGTGTGAAAGGAACTGAGTGAAGAACTGGAACCTGGACAGAGAGCTCT	1620
Db	1492	AGTGGACAGGCGAGTGTGAAAGGAACTGAGTGAAGAACTGGAACCTGGACAGAGAGCTCT	1551
QY	1621	GGCTTCCAAACAGCTGCAAAATGATGAAATGAAGCAAAACCAATTCGCAAGCAGGAAGGA	1680
Db	1552	GGCTTCCAAACAGCTGCAAAATGATGAAATGAAGCAAAACCAATTCGCAAGCAGGAAGGA	1611
QY	1681	CCTGGAACCATGACCATCTCAGGGCTCAGATGGAAGTTACTGTTCTGATTTTCATGC	1740
Db	1612	CCTGGAACCATGACCATCTCAGGGCTCAGATGGAAGTTACTGTTCTGATTTTCATGC	1671
QY	1741	TGAAAGCAGCAGCAGAGAGAAAAATTCATCAGGAAAAAGGAGCAACTGGCATTCGAGCTGGC	1800
Db	1672	TGAAAGCAGCAGCAGAGAGAAAAATTCATCAGGAAAAAGGAGCAACTGGCATTCGAGCTGGC	1731
QY	1801	AGTTCTGCTGAAAGAGATGATGCTTTTCGAAGACGAGGCGAGGCTCCTTTGATGAGAT	1860
Db	1732	AGTTCTGCTGAAAGAGATGATGCTTTTCGAAGACGAGGCGAGGCTCCTTTGATGAGAT	1791
QY	1861	GCAGATGCTGATGCGGCGCAGAAACAGTGAATCTGACAGAGGCTTACCTGTTCAAAG	1920
Db	1792	GCAGATGCTGATGCGGCGCAGAAACAGTGAATCTGACAGAGGCTTACCTGTTCAAAG	1851
QY	1921	AGGAGCTGAGGACAGGAGCTGGCGCAACAGCGGAATATTCGATTCATTCCTGCCCAA	1980
Db	1852	AGGAGCTGAGGACAGGAGCTGGCGCAACAGCGGAATATTCGATTCATTCCTGCCCAA	1911
QY	1981	GTGTGGAGAGGTTCTCCCTGACATAGACAGCTTACAGATTCACGTGATGGAATGCAATCAT	2040
Db	1912	GTGTGGAGAGGTTCTCCCTGACATAGACAGCTTACAGATTCACGTGATGGAATGCAATCAT	1971
QY	2041	TTAAGTGTGATGATACCTCCCAAAAACCTGTTGGT	2077
Db	1972	TTAAGTGTGATGATACCTCCCAAAAACCTGTTGGT	2008

ID ACF03997 standard; cDNA; 1856 BP.
XX AC ACF03997;
XX DT 23-SEP-2003 (first entry)
XX DE Human optineurin isoform 2 encoding cDNA SEQ ID NO:3.
XX KW Human; optineurin; OPTN; chromosome 10; 10p14; glaucoma; ophthalmic;
KW gene therapy; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 90..1823
FT FT /*tag= a
FT FT /product= "optineurin isoform 2"
PN W02003056037-A1.
XX 10-JUL-2003.
XX 18-DEC-2002; 2002WO-US041116.
XX 24-DEC-2001; 2001US-0344754P.
PR 30-JAN-2002; 2002US-00060981.
PR 28-FEB-2002; 2002US-00090118.
PR 25-OCT-2002; 2002US-00281454.
XX (UYCO-) UNIV CONNECTICUT.
PA (SGEO-) ST GEORGES ENTERPRISES LTD.
XX Sarfarazi M, Rezaie T, Child AH;
XX WPI: 2003-598273/56.
DR P-PSDB; ABR82078.
XX Diagnosing the presence or absence of optineurin-associated glaucoma or
PT optineurin-associated increased risk of glaucoma in an individual by
PT assessing a test sample for the presence or absence of an alteration in
PT the optineurin gene.
XX Claim 1; Page 72-75; 83pp; English.
XX The present invention describes a method (M1) for diagnosing the presence
CC or absence of optineurin-associated glaucoma or an optineurin-associated
CC increased risk of glaucoma in an individual comprising detecting the
CC presence or absence of an alteration in expression, composition or
CC activity of an optineurin nucleic acid or polypeptide, which is
CC indicative of the presence or absence, respectively, of the optineurin-
CC associated glaucoma. Optineurin has ophthalmic activity and can be used
CC in gene therapy. Optineurin therapeutic agents can be used for the
CC manufacture of a medicament for the treatment of glaucoma or of an
CC increased risk for glaucoma. Human optineurin is located on chromosome
CC 10, more specifically to 10p14. The present sequence encodes human
CC optineurin isoform 2, which is used in the exemplification of the present
CC invention
XX Sequence 1856 BP; 625 A; 382 C; 509 G; 340 T; 0 U; 0 Other;
Query Match 85.7%; Score 1779.4; DB 9; Length 1856;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1780; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 297 CAGGAATTCGCAATGCCATCAACCTCTCAGCTGCCTCCTGAAAGAGGACAGCC 356
DB 76 CAGGAATTCGCAATGCCATCAACCTCTCAGCTGCCTCCTGAAAGAGGACAGCC 135
QY 357 CCAGTGAAGACACAGGAAATGGACCCGCCACCTGGCCACCCAAACCTGGACACGTTTA 416
DB 136 CCAGTGAAGACACAGGAAATGGACCCGCCACCTGGCCACCCAAACCTGGACACGTTTA 195
QY 417 CCCGGAGGAGCTGCTGACGACGATGAAGAGCTCCTGACCAAGAACACCACGCTGAAG 476

196 CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCCTGACCGAGAACCCACGAGCTGAAG 255
QY 477 AAGCCATGAAGCTTAATTAATCAAGCCATGAAGGAGATTGTGAGAGCTTTGGCGCTGGA 536
DB 256 AAGCCATGAAGCTTAATTAATCAAGCCATGAAGGAGATTGTGAGAGCTTTGGCGCTGGA 315
QY 537 CAGAGAACAAGAGAAAGCCAGCTTTTGTAGATACAGAGCAAGAGCAAAAGAGC 596
DB 316 CAGAGAACAAGAGAAAGCCAGCTTTTGTAGATACAGAGCAAGAGCAAAAGAGC 375
QY 597 GTCTAATGGCTTGTAGTCATGAGAAATGAGAAATTTGAAGAGAGCTTGGAAAACTAAAG 656
DB 376 GTCTAATGGCTTGTAGTCATGAGAAATGAGAAATTTGAAGAGAGCTTGGAAAACTAAAG 435
QY 657 GGAATCAGAAAGCTCATCTGAGGACCCACTGATGACTCCAGGCTTCCAGGGCCGAG 716
DB 436 GGAATCAGAAAGCTCATCTGAGGACCCACTGATGACTCCAGGCTTCCAGGGCCGAG 495
QY 717 CGGAGCAGGAAAGGACCCAGCTCAGGACCCAGGCTGAGGCTACAAGCAGAGAGGCGAG 776
DB 496 CGGAGCAGGAAAGGACCCAGCTCAGGACCCAGGCTGAGGCTACAAGCAGAGAGGCGAG 555
QY 777 ACCTGTTGGCATCGTGTCTGAACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAG 836
DB 556 ACCTGTTGGCATCGTGTCTGAACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAG 615
QY 837 ATTCTTTTGTGAAATTTAGGATGGCTGAAGAGAGCAGAGGCTGATGAAAGAAATCA 896
DB 616 ATTCTTTTGTGAAATTTAGGATGGCTGAAGAGAGCAGAGGCTGATGAAAGAAATCA 675
QY 897 AGCATAGTCTGGGCCCCAGAGAACAGCTCTCCATGGCACGGCATCTCTAAATATAGGA 956
DB 676 AGCATAGTCTGGGCCCCAGAGAACAGCTCTCCATGGCACGGCATCTCTAAATATAGGA 735
QY 957 GCAGATCTGCAGATGGGGCCCAAGAAATTACTTCAACATGAGGAGTTAACTGTGAGCCAGC 1016
DB 736 GCAGATCTGCAGATGGGGCCCAAGAAATTACTTCAACATGAGGAGTTAACTGTGAGCCAGC 795
QY 1017 TCCTGCTGCTCTAAGGAGGGAATCAGAGGTGGAGAGACTTGAAGTTGCACCTCAGG 1076
DB 796 TCCTGCTGCTCTAAGGAGGGAATCAGAGGTGGAGAGACTTGAAGTTGCACCTCAGG 855
QY 1077 AGGCCAAAGAAAGAGTTTCAAGATTTTGAAGAAAGAAACAAGTAATCGTTCTGAGATTGAAA 1136
DB 856 AGGCCAAAGAAAGAGTTTCAAGATTTTGAAGAAAGAAACAAGTAATCGTTCTGAGATTGAAA 915
QY 1137 CCAGACAGAGGGGAGCAGAGAGAAATGATGAAGAGAAAGGCCCGGAGACTCTGTTG 1196
DB 916 CCAGACAGAGGGGAGCAGAGAGAAATGATGAAGAGAAAGGCCCGGAGACTCTGTTG 975
QY 1197 GAAGCGAAGTGGAAAGCACTGAACCTCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGG 1256
DB 976 GAAGCGAAGTGGAAAGCACTGAACCTCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGG 1035
QY 1257 CTCATCAAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAG 1316
DB 1036 CTCATCAAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAG 1095
QY 1317 CCCTTGAAAGGAAAAATTTCTGCAATTTCCATCAGAGTTGAATGAAAAAGCAGAGCTGTTTT 1376
DB 1096 CCCTTGAAAGGAAAAATTTCTGCAATTTCCATCAGAGTTGAATGAAAAAGCAGAGCTGTTTT 1155
QY 1377 ATACTAACAAAAAGTTAGAGCTACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAAC 1436
DB 1156 ATACTAACAAAAAGTTAGAGCTACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAAC 1215
QY 1437 AGGCTAAACAGAGGATGAAGTCCAAATTAATCTGCTACAGATGACACACAAAGC 1496
DB 1216 AGGCTAAACAGAGGATGAAGTCCAAATTAATCTGCTACAGATGACACACAAAGC 1275
QY 1497 TTCTTCAAGAACATAATAATGCAATTGAAAAACAATTGAGGAACCTTAAGAAAAAGAGTCAAG 1556

DR	P-PSDB; ADO43186. GENBANK; AF420372.	DR	1276	TTCTTTCAAGAACATAATATGCAATGTGAAAACAAATTGAGGAACTAAACAAGAAAAAGAGTCAG	1335
DR		DR	1557	AAAAAGTGGACAGGCGAGTGCTGGAAGGAACCTGAGTGAAGAAAACCTGGAATCTGGCAGAGAAGG	1616
PT	New optineurin nucleic acids and polypeptides having sequence alterations, which indicates the presence of an optineurin-associated (risk of) glaucoma, useful for detecting or diagnosing glaucoma, e.g. primary open angle glaucoma.	PT	1336	AAAAAGTGGACAGGCGAGTGCTGGAAGGAACCTGAGTGAAGAAAACCTGGAATCTGGCAGAGAAGG	1395
PT		PT	1617	CTCTGGCTTCCAAAACAGCTGCAAAATGGAATGGAATGGAAGCAAAACCAATTGCCAAGCAGGAAG	1676
XX		XX	1396	CTCTGGCTTCCAAAACAGCTGCAAAATGGAATGGAATGGAAGCAAAACCAATTGCCAAGCAGGAAG	1455
PS	Claim 1; SEQ ID NO 3; 110pp; English.	PS	1677	AGGACCTGGAACCAATGACCATCTCCAGGCTCAGATGGAATTTACTGTCTGATTTTC	1736
CC	The present sequence is that of cDNA encoding isoform 2 of human optineurin (optic neuropathy inducing protein). Mutation of the optineurin gene on chromosome 10 is associated with primary open angle glaucoma. In specific embodiments of the invention, the alteration is (numbering according to optineurin isoform 1 cDNA ADO43183): a change from GAG to AAG at codon 50 of the optineurin gene; an insertion of AG after codon 127; a change from CGG to CAG at codon 545; or a combination of one or more of these. These alterations are associated with glaucoma, and the presence of one or more of these alterations is diagnostic for glaucoma. In another embodiment, the alteration is a change from AAG to AAG at codon 98. The presence of this alteration is indicative of an increased risk of glaucoma, and is diagnostic of this increased risk.	CC	1456	AGGACCTGGAACCAATGACCATCTCCAGGCTCAGATGGAATTTACTGTCTGATTTTC	1515
CC	Other alterations include a change from CCC to GCC at codon 16, a change from CAG to CAC at codon 42, a change from GAA to GTA at codon 92, and a change from GAA to AAA at codon 322. Alterations include combinations of alterations such as combinations of alterations associated with glaucoma and those associated with an increased risk of glaucoma. Methods of detection, prognosis and diagnosis of the presence or absence of optineurin-associated glaucoma or of an optineurin-associated increased risk of glaucoma are described, in which a sample is tested for the presence of these gene sequence alterations or for alterations in the expression or activity of the optineurin protein. Also described are methods of therapy of glaucoma, including gene therapy methods.	CC	1737	ATGCTGAAAGACGACGAGAGAGAAAAATTCATGAGGAAAAGGACCACTGGCATTCGCAGC	1796
CC		CC	1516	ATGCTGAAAGACGACGAGAGAAAAATTCATGAGGAAAAGGACCACTGGCATTCGCAGC	1575
CC		CC	1797	TGGCAGTCTTCTGAAAGAGAAATGATGCTTTTTCGAAGCGGAGGACGAGCTCTTGATGG	1856
CC		CC	1576	TGGCAGTCTTCTGAAAGAGAAATGATGCTTTTTCGAAGCGGAGGACGAGCTCTTGATGG	1635
CC		CC	1857	AGATGACAGTCTGATCGGGCGGAGAAACAAGTGAATCTGACACAGGCTTACCTTGTTTC	1916
CC		CC	1636	AGATGACAGTCTGATCGGGCGGAGAAACAAGTGAATCTGACACAGGCTTACCTTGTTTC	1695
XX		XX	1917	AAAGAGGAGCTGAGGACAGGAGCTGCGGCAACAGCGGAATATTCGGATTTCCTGACC	1976
SQ	Sequence 1856 BP; 625 A; 382 C; 509 G; 340 T; 0 U; 0 Other;	SQ	1696	AAAGAGGAGCTGAGGACAGGAGCTGCGGCAACAGCGGAATATTCGGATTTCCTGACC	1755
			1977	CCAAGTGTGAGAGGTTCTGCCTGACATAGACACGCTTACAGATTACGATGATTTGCA	2036
			1756	CCBAGTGTGAGAGGTTCTGCCTGACATAGACACGCTTACAGATTACGATGATTTGCA	1815
			2037	TCATTTAAGTGTGATGATATCACTCCGCCAAAACCTGTTGGT	2077
			1816	TCATTTAAGTGTGATGATATCACTCCGCCAAAACCTGTTGGT	1856
			RESULT 13		
			ADO43185		
			ID	ADO43185 standard; cDNA; 1856 BP.	
			XX		
			AC	ADO43185;	
			XX		
			DT	29-JUL-2004 (first entry)	
			XX		
			DE	Human optineurin isoform 2 cDNA, associated with glaucoma.	
			XX		
			KW	Human; optineurin; glaucoma; diagnosis; gene therapy; ophthalmological;	
			XX	gene; chromosome 10p14; ss.	
			XX		
			OS	Homo sapiens.	
			XX		
			XX		
			FT	Key	
			CD	Location/Qualifiers	
			XX	90..1823	
			FT	/*tag= a	
			XX	/product= "Human optineurin isoform 1"	
			XX		
			XX	WO2004039312-A2.	
			PN		
			PD	13-MAY-2004.	
			XX		
			PF	26-JUN-2003; 2003WO-US020165.	
			XX		
			PR	25-OCT-2002; 2002US-00281457.	
			XX		
			PA	(UYCO-) UNIV CONNECTICUT	
			XX	(SGEO-) ST GEORGES ENTERPRISES LTD.	
			XX		
			PI	Sarfarazi M, Rezaie T, Child AH;	
			XX		
			XX	WPI; 2004-376046/35.	
			DR		

Db 556 ACCTGTTGGGATCGTGTCTGAACCTGACGCTCAAGCTGAACTCCAGGGCTCCTCAGAAG 615
Qy 837 ATTCCTTTGTTGAAATAGGATGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA 896
Db 616 ATTCCTTTGTTGAAATAGGATGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA 675
Qy 897 AGCATAGTCTGGGCCCCACGAGAAACAGTCTCCACTGCGACGGCAATTCGTAATAATAGGA 956
Db 676 AGCATAGTCTGGGCCCCACGAGAAACAGTCTCCACTGCGACGGCAATTCGTAATAATAGGA 735
Qy 957 GCAGATCTGCAGATGGGGCCCAAGAAATACCTCGAACATGAGGATTAACCTGTGAGCCAGC 1016
Db 736 GCAGATCTGCAGATGGGGCCCAAGAAATACCTCGAACATGAGGATTAACCTGTGAGCCAGC 795
Qy 1017 TCCTGCTGTGCTTAAGGGAAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGG 1076
Db 796 TCCTGCTGTGCTTAAGGGAAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGG 855
Qy 1077 AGGCCAAAGAAAGATTTTCAGATTTTGAAGAAACAAAGTAATCGTCTGAGATTGAAA 1136
Db 856 AGGCCAAAGAAAGATTTTCAGATTTTGAAGAAACAAAGTAATCGTCTGAGATTGAAA 915
Qy 1137 CCCAGACAGAGGGGACACAGAGAAAGAGATGATGAAGAAAGGCCCGGAGACTGTTG 1196
Db 916 CCCAGACAGAGGGGACACAGAGAAAGAGATGATGAAGAAAGGCCCGGAGACTGTTG 975
Qy 1197 GAAGCGAAGTGGAGACCTGAACCTCCAGGTGACATCTCTGTTTAAAGGACTTCAAGAGG 1256
Db 976 GAAGCGAAGTGGAGACCTGAACCTCCAGGTGACATCTCTGTTTAAAGGACTTCAAGAGG 1035
Qy 1257 CTCATCAAAAACCTCAGCGAAGCTGAGCTAATGAAGAGAGAGACTTCAAGAAAAGTGTGAGG 1316
Db 1036 CTCATCAAAAACCTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTGAGG 1095
Qy 1317 CCCTTGAAGGAAAAATCTGCAATTCCTCAGATGAGTTGAATGAAGAAAGAGAGCTGTTT 1376
Db 1096 CCCTTGAAGGAAAAATCTGCAATTCCTCAGATGAGTTGAATGAAGAAAGAGAGCTGTTT 1155
Qy 1377 ATACTTAACAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC 1436
Db 1156 ATACTTAACAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC 1215
Qy 1437 AGGCTAAAAACAGAGGATGAAAAGTCCAAATTAACCTGTCTACAGATGACACACAAACAGC 1496
Db 1216 AGGCTAAAAACAGAGGATGAAAAGTCCAAATTAACCTGTCTACAGATGACACACAAACAGC 1275
Qy 1497 TTCTTCAAGAAACATAATATGCAATTTGAAAACAATTTGAGGAACTAAACAAGAAAAAGAGTCAG 1556
Db 1276 TTCTTCAAGAAACATAATATGCAATTTGAAAACAATTTGAGGAACTAAACAAGAAAAAGAGTCAG 1335
Qy 1557 AAAAAGTGGACGGGAGTCTGAAGAACTGAGTGAAGAACTGGAACCTGGACCTGGCAGAGAGG 1616
Db 1336 AAAAAGTGGACGGGAGTCTGAAGAACTGAGTGAAGAACTGGAACCTGGACCTGGCAGAGAGG 1395
Qy 1617 CTCTGGCTTCCAAACAGCTCAAAATGATGAATGAAGCAAAACCAATTCGCAACGAGGAAG 1676
Db 1396 CTCTGGCTTCCAAACAGCTCAAAATGATGAATGAAGCAAAACCAATTCGCAACGAGGAAG 1455
Qy 1677 AGACCTGGAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTC 1736
Db 1456 AGACCTGGAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTC 1515
Qy 1737 ATGCTGAAGAGCAGCGAGAGAGAAATTCATCAGGAAAGGAGCACTGGCAATTCGAGC 1796
Db 1516 ATGCTGAAGAGCAGCGAGAGAGAAATTCATCAGGAAAGGAGCACTGGCAATTCGAGC 1575
Qy 1797 TGGCAGTCTCTGCTCAAGAGAAATGATGCTTTTCAAGACGGAGCGAGCAGCTCTCTGATGG 1856
Db 1576 TGGCAGTCTCTGCTCAAGAGAAATGATGCTTTTCAAGACGGAGCGAGCAGCTCTCTGATGG 1635
Qy 1857 AGATGAGAGTCTCTATGGGGCGAGAACAGTCACTCTGACCGACGAGCTTACCTTGTTC 1916

Db 1636 AGATGCAGAGTCGTCTATGGGGCGAGAACAAAGTGACTCTGACCAGCAGGCTTACCTTGTTC 1695
Qy 1917 AAAGAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAAATATTCGATTCATTCCTGCC 1976
Db 1696 AAAGAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAAATATTCGATTCATTCCTGCC 1755
Qy 1977 CCAAGTGTGAGAGGTTCTGCTTGCATACATACACAGCTTACAGATTCACGTCATGGATTGCA 2036
Db 1756 CCAAGTGTGAGAGGTTCTGCTTGCATACATACACAGCTTACAGATTCACGTCATGGATTGCA 1815
Qy 2037 TCATTTAAGTGTGATGATGATACCTCCCAAAACTGTTGTT 2077
Db 1816 TCATTTAAGTGTGATGATGATACCTCCCAAAACTGTTGTT 1856

RESULT 14
ADR14212
ID ADR14212 standard; DNA; 2327 BP.
XX
AC ADR14212;
XX
XX 21-OCT-2004 (first entry)
XX
XX Human NF-kappaB pathway-associated gene SeqID213.
DE
XX
XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
KW antiarthritic; antiirheumatic; gastrointestinal-Gen; antiasthmatic;
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW immunosuppressive; vulnery; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW hyper-IGM syndrome; hypohidrotic ectodermal dysplasia;
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;
KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; gene; ds; human.
XX
OS Homo sapiens.
XX
XX WO2004065577-A2.
XX
XX 05-AUG-2004.
XX
XX 13-JAN-2004; 2004WO-US000798.
XX
XX 14-JAN-2003; 2003US-0440068P.
XX 12-MAY-2003; 2003US-0469757P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Nadler SG, Neubauer MG, Feder JN, Carman J;
XX
XX WPI; 2004-562168/54.
XX P-PSDB; ADR14213.
XX
XX New isolated polynucleotides and polypeptides associated with NF-kappaB
XX pathway, useful for diagnosing, treating, or preventing disorders or
XX diseases associated with NF-kappaB pathway.
XX
XX Claim 1; SEQ ID NO 213; 237pp; English.
XX
XX This invention relates to the novel association of protein sequences (and
XX the genes which encode them) to the NF-kappaB pathway. The invention may
XX be useful for the production of compounds with an antiinflammatory,
XX cytostatic, hepatotropic, virucide, antiarthritic, antiirheumatic,
XX gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,
XX immunomodulator, cerebroprotective, vasotropic, immunosuppressive or

CC vulnery activity or for gene therapy. The proteins and nucleotides are
CC useful for diagnosing, preventing, treating, or ameliorating conditions
CC or diseases associated with the NF-kappaB pathway. The condition is an
CC immune disorder, an inflammatory disorder, an inflammatory disorder
CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC immune activity, disorders related to aberrant acute phase responses,
CC hypergenital conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, conditions related to organ transplant
CC rejection, disorders related to aberrant signal transduction,
CC proliferating disorders, cancers and HIV propagation in cells infected
CC with other viruses. The present sequence is that of a human gene which is
CC subject to the novel association with the NF-kappaB pathway of the
CC invention. Note: This sequence does not appear in the specification but
CC was obtained by the indexer from Genbank.

XX
SQ Sequence 2327 BP; 747 A; 499 C; 636 G; 445 T; 0 U; 0 Other;

Query Match 85.7%; Score 1779.4; DB 13; Length 2327;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1780; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	297	CAGGAACCTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGGGAGCAGCC	356
DB	141	CAGGAACCTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGGGAGCAGCC	200
QY	357	CCAGTGAAGACACAGGAATGGACCCGCCACCTGGCCCAACCCAAACCTGGACACCTTTA	416
DB	201	CCAGTGAAGACACAGGAATGGACCCGCCACCTGGCCCAACCCAAACCTGGACACCTTTA	260
QY	417	CCCCGGAGGAGCTGCTGCAGCAGATGAAGAGGCTCTGACCAAGAACCCACGCTGAAAG	476
DB	261	CCCCGGAGGAGCTGCTGCAGCAGATGAAGAGGCTCTGACCGAGAACCCACGCTGAAAG	320
QY	477	AAGCCATGAGCTTAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTCGGCTTGA	536
DB	321	AAGCCATGAGCTTAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTCGGCTTGA	380
QY	537	CAGAGAAACAGAGAGGAGAACCGCAGTTTTTTTGAGATACAGAGCAAGAGCAAAAGAGC	596
DB	381	CAGAGAAACAGAGAGGAGAACCGCAGTTTTTTTGAGATACAGAGCAAGAGCAAAAGAGC	440
QY	597	GTCTAATGGCTTGAATGAGAAATGAGAAATTTGAAGGAAGAGCTTGGAAACTTAAAG	656
DB	441	GTCTAATGGCTTGAATGAGAAATGAGAAATTTGAAGGAAGAGCTTGGAAACTTAAAG	500
QY	657	GGAAATCAGAAAGGTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAG	716
DB	501	GGAAATCAGAAAGGTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAG	560
QY	717	CGAGCAGGAAAGGACAGCTCAGGACCCAGGTGGTGGGCTACAGCAGAGAGGCGAG	776
DB	561	CGAGCAGGAAAGGACAGCTCAGGACCCAGGTGGTGGGCTACAGCAGAGAGGCGAG	620
QY	777	ACCTGTTGGGCATCGTGTCTGAACTGCAGTCTCAAGCTGAACTCAGCGGCTCTCAGAAG	836
DB	621	ACCTGTTGGGCATCGTGTCTGAACTGCAGTCTCAAGCTGAACTCAGCGGCTCTCAGAAG	680
QY	837	ATTCTTTTGTGAAATTAGATGGCTCAAGGAGAGCAGAGGCTCAGTAAAGAAATCA	896
DB	681	ATTCTTTTGTGAAATTAGATGGCTCAAGGAGAGCAGAGGCTCAGTAAAGAAATCA	740
QY	897	AGCATAGCTCTGGGCCCAACAGAAACAGTCTCCACTGGCAGGCACTTCTTAAATATAGGA	956
DB	741	AGCATAGCTCTGGGCCCAACAGAAACAGTCTCCACTGGCAGGCACTTCTTAAATATAGGA	800
QY	957	GCAGATCTGCAGATGGGGCCCAAGAAATTAATCTTCGAACTAGAGGATTAATCTGTGAGCCAGC	1016

DB	801	GCAGATCTGCAGATGGGGCCCAAGAAATTAATCTTCGAACTAGAGGATTAATCTGTGAGCCAGC	860
QY	1017	TCCTCTCTGCTCCTAAGAGGAGGGAATCAGAAGGTGGAGAGACTTGAAGTTGCACCTCAAG	1076
DB	861	TCCTCTCTGCTCCTAAGAGGAGGGAATCAGAAGGTGGAGAGACTTGAAGTTGCACCTCAAG	920
QY	1077	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAA	1136
DB	921	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAA	980
QY	1137	CCGACACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACAGTTGTG	1196
DB	981	CCGACACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACAGTTGTG	1040
QY	1197	GAAGCGAAGTGAAGACACTGAAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAG	1256
DB	1041	GAAGCGAAGTGAAGACACTGAAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAG	1100
QY	1257	CTCATCAAAACTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAGTGTCAAG	1316
DB	1101	CTCATCAAAACTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAGTGTCAAG	1160
QY	1317	CCCTTGAAGGAAATTTCTGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTGTT	1376
DB	1161	CCCTTGAAGGAAATTTCTGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTGTT	1220
QY	1377	ATACTAAACAAAAGTTAGAGCTACAAAGTGAAGCATGCTATCAGAAATCAAAATGGAAC	1436
DB	1221	ATACTAAACAAAAGTTAGAGCTACAAAGTGAAGCATGCTATCAGAAATCAAAATGGAAC	1280
QY	1437	AGGCTAAACACAGAGATGAAAAGTCCAAATTAATGTTGTTGCTACAGATGACACAAAGC	1496
DB	1281	AGGCTAAACACAGAGATGAAAAGTCCAAATTAATGTTGTTGCTACAGATGACACAAAGC	1340
QY	1497	TTCTTTCAGAAACATTAATGATGATTAAGAACTTGAAGCACTTAAACAGAAAGAGTCTAG	1556
DB	1341	TTCTTTCAGAAACATTAATGATGATTAAGAACTTGAAGCACTTAAACAGAAAGAGTCTAG	1400
QY	1557	AAAAAGTGAACAGGGCAGTGTCTGAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAG	1616
DB	1401	AAAAAGTGAACAGGGCAGTGTCTGAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAG	1460
QY	1617	CTCTGCTTCCAAAACAGCTGCAAAATGGAATGAAGCAAAACCTTCCAGCAGAGAG	1676
DB	1461	CTCTGCTTCCAAAACAGCTGCAAAATGGAATGAAGCAAAACCTTCCAGCAGAGAG	1520
QY	1677	AGGACCTGAAACCACTGACCATCTCAGGGCTCAGATGAAAGTTTACTGTTCTGATTTTC	1736
DB	1521	AGGACCTGAAACCACTGACCATCTCAGGGCTCAGATGAAAGTTTACTGTTCTGATTTTC	1580
QY	1737	ATGCTGAAAGAGCAGCGAGAGAGAAATTCATGAGAAAGAGCAACTGGCAATTCGAGC	1796
DB	1581	ATGCTGAAAGAGCAGCGAGAGAGAAATTCATGAGAAAGAGCAACTGGCAATTCGAGC	1640
QY	1797	TGCGAGTTCTGCTGAAGAGAAATGATGCTTTTGAAGACCGAGGCGAGCTTCTGATGG	1856
DB	1641	TGCGAGTTCTGCTGAAGAGAAATGATGCTTTTGAAGACCGAGGCGAGCTTCTGATGG	1700
QY	1857	AGATGAGAGTCTGATGAGGCGAGAACTGACCTCTGACAGCAGGCTTACTTGTGTTTC	1916
DB	1701	AGATGAGAGTCTGATGAGGCGAGAACTGACCTCTGACAGCAGGCTTACTTGTGTTTC	1760
QY	1917	AAAGAGGAGCTGAGGACAGGGAATGCGCGCAACAGCGGAAATATTCGATTCATTCCTGCC	1976
DB	1761	AAAGAGGAGCTGAGGACAGGGAATGCGCGCAACAGCGGAAATATTCGATTCATTCCTGCC	1820
QY	1977	CAAAGTGTGAGAGGTTTCTGCTGACATGACAGTTACAGTTACAGTTACAGTTGATGCA	2036
DB	1821	CAAAGTGTGAGAGGTTTCTGCTGACATGACAGTTACAGTTACAGTTACAGTTGATGCA	1880
QY	2037	TCATTTTAAGTTGATGATATCACTCCCAAAACTGTTGGT	2077

Db 1881 TCATTAAAGTTGATGATACACCTCCCAAAACATGTTGGT 1921
RESULT 15
ADP23351
XX ADP23351 standard; cDNA; 2327 BP.
XX
AC ADP23351;
XX
XX 18-NOV-2004 (first entry)
XX
XX PRO polypeptide encoding cDNA SEQ ID NO:445.
XX
XX ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
XX Unidentified.
XX
XX WO2004041170-A2.
XX
XX 21-MAY-2004.
XX
XX 30-OCT-2003; 2003WO-US034312.
XX
XX 01-NOV-2002; 2002US-0423394P.
XX
XX (GETH) GENENTECH INC.
XX
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;
PI Wu TD;
XX
XX WPI; 2004-419628/39.
XX
XX P-PSDB; ADP23352.
XX
XX New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
XX Claim 1; SEQ ID NO 445; 2940pp; English.
XX
XX The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, an
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
XX the invention.
XX
XX Sequence 2327 BP; 747 A; 499 C; 636 G; 445 T; 0 U; 0 Other;

	Query Match	85.7%	Score 1779.4;	DB 13;	Length 2327;
	Best Local Similarity	99.9%	Pred. No. 0;		
	Matches 1780;	Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;
Qy	297	CAGGAACCTTCTGCATATGTCATCAACCTCTCAGCTGCTCTCAGTGAAGAGGACAGCC	356		
Db	141	CAGGAACCTTCTGCATATGTCATCAACCTCTCAGCTGCTCTCAGTGAAGAGGACAGCC	200		
Qy	357	CCAGTGAAGCAGCAGGAATGGACCCCTCCACCTGGCCCTCCCAAACTTGACAGCTTTA	416		
Db	201	CCAGTGAAGCAGCAGGAATGGACCCCTCCACCTGGCCCTCCCAAACTTGACAGCTTTA	260		
Qy	417	CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTTACCAGAAACACACAGCTGAAG	476		
Db	261	CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTTACCAGAAACACACAGCTGAAG	320		
Qy	477	AAGCCATGAAGCTTAATAATCAAGCCATGAAGGAGATTTGAGGAGCTTTCCGCCCTGA	536		
Db	321	AAGCCATGAAGCTTAATAATCAAGCCATGAAGGAGATTTGAGGAGCTTTCCGCCCTGA	380		
Qy	537	CAGAGAAACAGAGGAAGACCGCAGCTTTTGGAGATACAGACCAAGAGCAAGAGAGC	596		
Db	381	CAGAGAAACAGAGGAAGACCGCAGCTTTTGGAGATACAGACCAAGAGCAAGAGAGC	440		
Qy	597	GTCTAATGGCTTTGAGTCATGAGAAATTTGAAGAGAGCTTTGAGAACTATAAG	656		
Db	441	GTCTAATGGCTTTGAGTCATGAGAAATTTGAAGAGAGCTTTGAGAACTATAAG	500		
Qy	657	GGAAATCAGAAAGTCTATCTGAGGACCCCTCATGATGATCTCAGGCTTCCAGGCCCAG	716		
Db	501	GGAAATCAGAAAGTCTATCTGAGGACCCCTCATGATGATCTCAGGCTTCCAGGCCCAG	560		
Qy	717	CGGAGCAGGAAAGGACCGCTCAGGACCCAGCTGAGGAGCTTGAAGAGAGCTTTGAGAA	776		
Db	561	CGGAGCAGGAAAGGACCGCTCAGGACCCAGCTGAGGAGCTTGAAGAGAGCTTTGAGAA	620		
Qy	777	ACCTGTTGGCATCGTGTCTGAACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAG	836		
Db	621	ACCTGTTGGCATCGTGTCTGAACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAG	680		
Qy	837	ATTCTCTTGTGTAATTAGGATCGCTGAGGAGACGAGAGAGGCTCAGTAAAGAAATCA	896		
Db	681	ATTCTCTTGTGTAATTAGGATCGCTGAGGAGACGAGAGAGGCTCAGTAAAGAAATCA	740		
Qy	897	AGCATAGTCTCTGGGCCACGAGAAACAGTCTCCACTGCGACGCGATCTCTAAATATAGGA	956		
Db	741	AGCATAGTCTCTGGGCCACGAGAAACAGTCTCCACTGCGACGCGATCTCTAAATATAGGA	800		
Qy	957	GCAGATCTGCAGATGGGGCCCAAGAAATTAATCTCGAAACATGAGGAGTTAACTGTGAGCCAGC	1016		
Db	801	GCAGATCTGCAGATGGGGCCCAAGAAATTAATCTCGAAACATGAGGAGTTAACTGTGAGCCAGC	860		
Qy	1017	TCTGCTGTGCTTAAGGGAAGGGAATCAGAAAGTGAGAGACTTTGAAGTTGCACTCAAGG	1076		
Db	861	TCTGCTGTGCTTAAGGGAAGGGAATCAGAAAGTGAGAGACTTTGAAGTTGCACTCAAGG	920		
Qy	1077	AGSCCAAGAAAGAGTTTCTGATTTTGAAGAAACAAAGTAATCGTCTGAGATTTGAAA	1136		
Db	921	AGSCCAAGAAAGAGTTTCTGATTTTGAAGAAACAAAGTAATCGTCTGAGATTTGAAA	980		
Qy	1137	CCGAGACAGAGGGGACACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTG	1196		
Db	981	CCGAGACAGAGGGGACACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTG	1040		
Qy	1197	GAAGCGAAGTGGAAGCACTGAACCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGG	1256		
Db	1041	GAAGCGAAGTGGAAGCACTGAACCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGG	1100		
Qy	1257	CTCATCAAACTCAGCGAAGCTGAGTAATGAAGAGAGACTTCAAGAAAGCTGTCAGG	1316		
Db	1101	CTCATCAAACTCAGCGAAGCTGAGTAATGAAGAGAGACTTCAAGAAAGCTGTCAGG	1160		
Qy	1317	CCCTTGAAAGGAAAAATTTCTGCAATTTCCATCAGAGTTGAATGAAGAGCAAGAGCTGTTT	1376		

Db 1161 ||||| CCCTTGAAGGAAAAATCTTGCATTTCCATCAGAGTTGAATGAAAAAGCAGGCTTGT 1220

Qy 1377 ATACTAAACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC 1436

Db 1221 ATACTAAACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC 1280

Qy 1437 AGSCTAAAAACAGAGGATGAAAAGTCCAAATTTAACTGTGTACAGATGACACACAACAGC 1496

Db 1281 AGSCTAAAAACAGAGGATGAAAAGTCCAAATTTAACTGTGTACAGATGACACACAACAGC 1340

Qy 1497 TTCTTCAAGAACATAATAATGCAATTGAAAACAAATGAGGAACCTAAACAAGAAAGATCAG 1556

Db 1341 TTCTTCAAGAACATAATAATGCAATTGAAAACAAATGAGGAACCTAAACAAGAAAGATCAG 1400

Qy 1557 AAAAGTGGACAGGCGAGTGTCTGAAGAACTGAGTGAATAAACTGGAACTGGCAGAGAAG 1616

Db 1401 AAAAGTGGACAGGCGAGTGTCTGAAGAACTGAGTGAATAAACTGGAACTGGCAGAGAAG 1460

Qy 1617 CTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAAG 1676

Db 1461 CTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAAG 1520

Qy 1677 AGSACTGGAACCATGACCATCTCAGGCTCAGATGAGNAGTTTACTGTTCTGATTTTC 1736

Db 1521 AGSACTGGAACCATGACCATCTCAGGCTCAGATGAGNAGTTTACTGTTCTGATTTTC 1580

Qy 1737 ATGCTGAAGAGCAGCGAGAGAGAAAATTCATGAGGAAGAGCAACTGGCATTGCAGC 1796

Db 1581 ATGCTGAAGAGCAGCGAGAGAGAAAATTCATGAGGAAGAGCAACTGGCATTGCAGC 1640

Qy 1797 TGGCAGTTCTGCTGAAAGAGAATGATGCTTTTTCGAAGACGAGGCGAGCTCCTTGATGG 1856

Db 1641 TGGCAGTTCTGCTGAAAGAGAATGATGCTTTTTCGAAGAGCGAGGCGAGCTCCTTGATGG 1700

Qy 1857 AGATGCAAGTCTGATGGGCGAGAACAACTGACTCTGACAGAGGCTTACTTGTTC 1916

Db 1701 AGATGCAAGTCTGATGGGCGAGAACAACTGACTCTGACAGAGGCTTACTTGTTC 1760

Qy 1917 AAAGAGAGCTGAGGACAGGAGCTGGCGCAACAGCGGAATATTCGGATTCATTCTGTGC 1976

Db 1761 AAAGAGAGCTGAGGACAGGAGCTGGCGCAACAGCGGAATATTCGGATTCATTCTGTGC 1820

Qy 1977 CCAAGTGTGAGAGGTTCTGCTGACATAGACAGCTTTACAGATTCAGTGATCGATTGCA 2036

Db 1821 CCAAGTGTGAGAGGTTCTGCTGACATAGACAGCTTTACAGATTCAGTGATCGATTGCA 1880

Qy 2037 TCATTTAAGTGTGATGATATCACTCCCAAACTGTTGGT 2077

Db 1881 TCATTTAAGTGTGATGATATCACTCCCAAACTGTTGGT 1921

RESULT 16

ADO43195

ID ADO43195 standard; cDNA; 2076 BP.

XX ADO43195;

AC ADO43195;

XX 29-JUL-2004 (first entry)

DT Macaque optineurin cDNA.

DE Mouse; optineurin; glaucoma; diagnosis; gene therapy; ophthalmological;

XX gene; ss.

XX Macaca mulatta.

XX Key Location/Qualifiers

FT CDS 241..1956

FT /*tag= a

FT /product= "Macaque optineurin"

XX

FN W02004039312-A2.

XX 13-MAY-2004.

XX 26-JUN-2003; 2003WO-US020165.

XX 25-OCT-2002; 2002US-00281457.

XX (UYCO-) UNIV CONNECTICUT.

PA (SGEO-) ST GEORGES ENTERPRISES LTD.

XX Sarfarazi M, Rezaie T, Child AH;

XX WPI; 2004-376046/35.

DR P-PSDB; ADO43196.

XX New optineurin nucleic acids and polypeptides having sequence

PT alterations, which indicates the presence of an optineurin-associated

PT (risk of) glaucoma, useful for detecting or diagnosing glaucoma, e.g.

PT primary open angle glaucoma.

XX Example; SEQ ID NO 13; 110pp; English.

XX The present sequence is that of macaque optineurin (optic neuropathy

CC inducing protein) cDNA. Comparison of protein sequences, and Western

CC blotting results, indicate conservation of optineurin as shown in human

CC ADO43184, macaque, mouse, rat, pig and cattle sequences. Overall, human

CC optineurin has 78-85% sequence identity with its homologues in mouse,

CC rat, pig and cattle, and 96% identity with macaque. Mutation of the human

CC optineurin gene is associated with primary open angle glaucoma. 2

CC Mutations, E50K and M98K, found in human optineurin are also found in

CC macaque, corroborating evidence that this mutation is a risk factor for

CC glaucoma. The E50K mutation is further conserved in mouse and bovine.

CC Methods of detection, prognosis and diagnosis of the presence or absence

CC of optineurin-associated glaucoma or of an optineurin-associated

CC increased risk of glaucoma are described, in which a sample is tested for

CC the presence of certain gene sequence alterations or for alterations in

CC the expression or activity of the optineurin protein. Also described are

CC methods of therapy of glaucoma, including gene therapy methods.

XX

SQ Sequence 2076 BP; 663 A; 443 C; 557 G; 413 T; 0 U; 0 Other;

Query Match 84.2%; Score 1747.8; DB 12; Length 2076;

Best Local Similarity 92.1%; Pred. No. 0;

Matches 1912; Conservative 0; Mismatches 77; Indels 88; Gaps 3;

Qy 1 ATCCCCGGTCGGAGTTCTCTCCAGGCGGCACGATGCCAGGAGCAGTAAACCTGAGCGA 60

Db 1 ATCCCCGGTCGGAGTTCTCTCCAGGCGGCACGATGCCAGGAGCAGTAAACCTGAGCGA 60

Qy 61 AGCCAAAGCCGGCGGCGAGGTGGCTTTGATAGCTGGTGGTGCACCTTCTGGCCCTTGA 120

Db 61 AGCCAAAGCCGGCGGCGAGGTGGCTTTGACACCTGGTGGTGCACCTTCTGGCTTGA 120

Qy 121 TGAGCCGTACGCTCTGTAAACCCCACTTCCTCACCTTTGAAACAGCTGCTGTTTCAGC 180

Db 121 TGAGCTGTACGCTCTGTAAACCCCACTTCCTCACCTAT-AAACAGCTGCTGTTTCAGC 179

Qy 181 ATTAATGAAGATTAGTCAAGTACAGCCCTGGTGTGCTGAGTCCGCACATAGAAGATCAA 240

Db 180 GTTAATGAAGATTAGTCAAGTACAGCCCTGGTGTGCTGAGTCCGCACATAGAAGATCAA 229

Qy 241 AAATGTCCAAATGTAACTGGAGAGAAAGTGGCAACTTTTGGAGTGACTTTTCCACAGG 300

Db 230 -----G 230

Qy 301 AACTTCTGCAATGTCCCATCAACCTCTCAGCTCCCTCAGTGAAGAGGAGCAGCCCGAG 360

Db 231 AACTTCTGCAATGTCCCATCAACCTCTCAGCTCCCTCAGTGAAGAGGAGCAGCCCGAG 290

Qy 361 TGAAGACACAGGAAATGGACCCCCCACTGGCCCAACCCCAACCTGAGACAGTTTACCCC 420

Db 291 TGAAGACACAGGAAATGGACCCCCCACTGGCCCAACCCCAACCTGAGACACATTACCCC 350


```
FT /frequency= "0.14"  
FT /note= "frequency in patient group (0.0% in non-patient  
FT group)"  
FT exon 627. .779  
FT /*tag= g  
FT /number= 8  
FT exon 780. .882  
FT /*tag= h  
FT /number= 9  
FT exon 883. .998  
FT /*tag= i  
FT /number= 10  
FT replace(898,a)  
FT /*tag= j  
FT /standard name= "Single nucleotide polymorphism"  
FT /frequency= "0.08"  
FT /note= "frequency in patient group (0.0% in non-patient  
FT group)"  
FT exon 999. .1148  
FT /*tag= k  
FT /number= 11  
FT exon 1149. .1242  
FT /*tag= l  
FT /number= 12  
FT exon 1243. .1401  
FT /*tag= m  
FT /number= 13  
FT exon 1402. .1532  
FT /*tag= n  
FT /number= 14  
FT exon 1533. .1612  
FT /*tag= o  
FT /number= 15  
FT exon 1613. .1734  
FT /*tag= p  
FT /number= 16  
XX  
XX EP1388590-A2.  
XX  
XX 11-FEB-2004.  
XX  
XX 29-JUL-2003; 2003EP-00447201.  
XX  
XX 02-AUG-2002; 2002JP-00226612.  
XX  
XX (SYSM-) SYSMEX CORP.  
XX  
XX Kouchi Y, Masago A, Takahata T;  
XX  
XX WPI; 2004-146134/15.  
XX  
XX Gene assay for predicting future onset of glaucoma, particularly primary  
XX open angle glaucoma or normal ocular tension glaucoma, comprises  
XX detecting a mutation of at least one base of the optineurin gene.  
XX  
XX Claim 2; SEQ ID NO 1; 31pp; English.  
XX  
XX The present sequence is the coding sequence of the glaucoma-associated  
XX gene, OPTN (optineurin). The invention relates to a gene assay method for  
XX predicting future onset of primary open angle glaucoma and/or normal  
XX ocular tension glaucoma. This involves detecting a mutation in the OPTN  
XX gene coding sequence, specifically a substitution of G for A at position  
XX 619 and/or a substitution of A for G at position 898 of the present  
XX sequence. The mutation(s) is detected using a nucleic acid amplification  
XX method.  
XX  
XX Sequence 1734 BP; 599 A; 347 C; 471 G; 317 T; 0 U; 0 Other;  
XX  
XX Query Match 83.3%; Score 1730.8; DB 12; Length 1734;  
XX Best Local Similarity 99.9%; Pred. No. 0;  
XX Matches 1732; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 311 ATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAGGAGGACAGCCCGCCAGTGAAGCACA 370
```

```
Db  
1 ATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAGGAGGACAGCCCGCCAGTGAAGCACA 60  
QY  
371 GGAATATGGAACCCCGCCACCTGGCCCAACCCAAACCTGGACACGTTTATCCCGGAGGAGCTG 430  
Db  
61 GGAATATGGAACCCCGCCACCTGGCCCAACCCAAACCTGGACACGTTTATCCCGGAGGAGCTG 120  
QY  
431 CTGACGACGATGAAAAGAGCTCTGACCAAGAACCAACCCAGCTGAAGAAGCCATGAAGCTA 490  
Db  
121 CTGACGACGATGAAAAGAGCTCTGACCGAGAACCAACCCAGCTGAAGAAGCCATGAAGCTA 180  
QY  
491 AATAATCAAGCCATGAAAAGGAGATTGAGGAGCTTTCCGCCCTGGACAGAGAAACAGAAG 550  
Db  
181 AATAATCAAGCCATGAAAAGGAGATTGAGGAGCTTTCCGCCCTGGACAGAGAAACAGAAG 240  
QY  
551 GAAGAACGCCAGTTTTTTTGAGATACAGAGCAAAAGAGCAAAAGAGCGTCTAATGGCCTTG 610  
Db  
241 GAAGAACGCCAGTTTTTTTGAGATACAGAGCAAAAGAGCAAAAGAGCGTCTAATGGCCTTG 300  
QY  
611 AGTCATGAGAAATGAGAAATTTGAAGAGAGCTTTGGAATACTAAAGGGGAATTCAGAAAGG 670  
Db  
301 AGTCATGAGAAATGAGAAATTTGAAGAGAGCTTTGGAATACTAAAGGGGAATTCAGAAAGG 360  
QY  
671 TCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGCGCGAAGCGGAGCAGGAAAAG 730  
Db  
361 TCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGCGCGAAGCGGAGCAGGAAAAG 420  
QY  
731 GACCACTCAGGACCCAGGTGGTGAAGCTACAAGCAGAGAAAGCAGACCTGTTGGGCATC 790  
Db  
421 GACCACTCAGGACCCAGGTGGTGAAGCTACAAGCAGAGAAAGCAGACCTGTTGGGCATC 480  
QY  
791 GTGTCTGAACTGAGCTCAAGCTGAACTCCAGGGGCTCTCAGAGAAATTCCTTTGTGAA 850  
Db  
481 GTGTCTGAACTGAGCTCAAGCTGAACTCCAGGGGCTCTCAGAGAAATTCCTTTGTGAA 540  
QY  
851 ATTAGGATGGCTGAAGGAGAAAGCAGAGGCTCAGTAAAAAGAAATCAAGCATAGTCTCTGGG 910  
Db  
541 ATTAGGATGGCTGAAGGAGAAAGCAGAGGCTCAGTAAAAAGAAATCAAGCATAGTCTCTGGG 600  
QY  
911 CCCACGAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGAGCAGATCTGCAGAT 970  
Db  
601 CCCACGAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGAGCAGATCTGCAGAT 660  
QY  
971 GGGGCCAAGAAATTAATTCGAAATGAGGAGTTAACTGTGAGCCAGCTCCTGCTGTCCTA 1030  
Db  
661 GGGGCCAAGAAATTAATTCGAAATGAGGAGTTAACTGTGAGCCAGCTCCTGCTGTCCTA 720  
QY  
1031 AGGGAAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAAGAAAGA 1090  
Db  
721 AGGGAAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAAGAAAGA 780  
QY  
1091 GTTTCAGATTTTGAAGAAAGAAACAAGTAATCGTTCTGAGATTGAAACCCAGAGAGGGG 1150  
Db  
781 GTTTCAGATTTTGAAGAAAGAAACAAGTAATCGTTCTGAGATTGAAACCCAGAGAGGGG 840  
QY  
1151 AGCAGAGAAAGAGAAATGATGAAGAAAGGCCCCCGAGACTGTTGGAGCCGAAGTGA 1210  
Db  
841 AGCAGAGAAAGAGAAATGATGAAGAAAGGCCCCCGAGACTGTTGGAGCCGAAGTGA 900  
QY  
1211 GCACTCAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAACTC 1270  
Db  
901 GCACTCAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAACTC 960  
QY  
1271 AGCGAAGCTGAGCTAAATGAAGAGAGACTTCAAGAAAAGTGTGAGGCCCTTTGAAGGAAA 1330  
Db  
961 AGCAAGCTGAGCTAAATGAAGAGAGACTTCAAGAAAAGTGTGAGGCCCTTTGAAGGAAA 1020  
QY  
1331 AATTCGCAATTCATCAGAGTTGAATGAAAAGCAGAGCTGTTTATATCTACAAAG 1390  
Db  
1021 AATTCGCAATTCATCAGAGTTGAATGAAAAGCAGAGCTGTTTATATCTACAAAG 1080  
QY  
1391 TTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAG 1450
```


Db	1081	TTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAACAGGCTTAAACACAGAG	1141
Qy	1451	GATGAAAGTCCAAATTAACATGTGCTACAGATGACACACAAACAGCTTCTTCAAGAACAT	1510
Db	1141	GATGAAAGTCCAAATTAACATGTGCTACAGATGACACACAAACAGCTTCTTCAAGAACAT	1200
Qy	1511	AATAATGCATTGAAAACAAATTTGAGGAACATAACAGAAAGAGTGCAGAAAAGTGGACAGG	1570
Db	1201	AATAATGCATTGAAAACAAATTTGAGGAACATAACAGAAAGAGTGCAGAAAAGTGGACAGG	1260
Qy	1571	GCAGTGTCTGAAGGAACCTGAGTGAATAAACTGGAACTGGCAGAGAAGGCTCTGGCTTCCAAA	1630
Db	1261	GCAGTGTCTGAAGGAACCTGAGTGAATAAACTGGAACTGGCAGAGAAGGCTCTGGCTTCCAAA	1320
Qy	1631	CAGCTGCAAAATGGAATGAAATGAACCAACCATTCGCCAAGCAGGAAGAGGACCTGGAAACC	1690
Db	1321	CAGCTGCAAAATGGAATGAAATGAACCAACCATTCGCCAAGCAGGAAGAGGACCTGGAAACC	1380
Qy	1691	ATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTCATGCTGAAAGACA	1750
Db	1381	ATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTCATGCTGAAAGACA	1440
Qy	1751	GCGAGACAGAAAAATTCAATGAGGAAAAAGGAGCAACTGGCATTTCAGCTGGCAGTTCTGCTG	1810
Db	1441	GCGAGACAGAAAAATTCAATGAGGAAAAAGGAGCAACTGGCATTTCAGCTGGCAGTTCTGCTG	1500
Qy	1811	AAAGAGAATGATGCTTTTCGAAAGCGAGGCGAGGCAGTCTTGTATGAGAGATGCAGAGTCGT	1870
Db	1501	AAAGAGAATGATGCTTTTCGAAAGCGAGGCGAGGCAGTCTTGTATGAGAGATGCAGAGTCGT	1560
Qy	1871	CATGGGCGAGAACCAAGTGACTCTGACCGAGGCTTACCTTGTTCGAAAGAGGAGCTGAG	1930
Db	1561	CATGGGCGAGAACCAAGTGACTCTGACCGAGGCTTACCTTGTTCGAAAGAGGAGCTGAG	1620
Qy	1931	GACAGGCACATGGCGGCAACAGCGGAATATTCGGATTCATTCTCGCCCCCAGATGTGGAG	1990
Db	1621	GACAGGCACATGGCGGCAACAGCGGAATATTCGGATTCATTCTCGCCCCCAGATGTGGAG	1680
Qy	1991	GTTCCTGCTGACATAGACAGCTTACAGATTTCACGTGATGATTGCATCATTTAA	2044
Db	1681	GTTCCTGCTGACATAGACAGCTTACAGATTTCACGTGATGATTGCATCATTTAA	1734
RESULT 18			
ADW86571			
ID	ADW86571 standard; DNA; 1734 BP.		
XX			
AC	ADW86571;		
XX			
DT	21-APR-2005 (first entry)		
XX			
DE	Human optineurin (OPTN) DNA amplified to indicate glaucoma onset Seq 1.		
XX			
KW	glaucoma; optineurin; ophthalmological; ds; ocular disease;		
KW	DNA amplification; genetic marker.		
XX			
OS	Homo sapiens.		
XX			
PH	Key		
FT	mutation		
FT	Location/Qualifiers		
FT	619		
FT	/*tag= a		
FT	/note= "Wild type A can be substituted by a G"		
FT	898		
FT	/*tag= b		
FT	/note= "Wild type G can be substituted by a A"		
XX			
PN	JP2005034112-A.		
XX			
PD	10-FEB-2005.		
XX			
PF	29-JUL-2003; 2003JP-00281897.		
XX			
PR	02-AUG-2002; 2002JP-00226612.		
XX			

30-JUN-2003; 2003JP-0018070.
(TOAI-) TOA IYO DENSHI KK.
WPI; 2005-156038/17.
Estimating risk of onset of glaucoma, involves analyzing mutation in any one portion of Optic new phosphorous gene, and utilizing analyzed mutation as index for estimating risk of onset of glaucoma.
Claim 2; SEQ ID NO 1; 13pp; Japanese.
This invention relates to a novel method for estimating the risk of onset of glaucoma. Specifically, it refers to a method that involves analyzing a mutation in any one portion of an optineurin (OPTN) gene and utilizing the analyzed mutation as an index for estimating the risk of onset of glaucoma. The present invention describes oligonucleotides to detect mutations that hybridize with one or more portions of the OPTN glaucoma related gene. Accordingly, it provides oligos that detect A619G and G898A mutations in the human OPTN gene of patients at risk of developing glaucoma, in particular primary open-angle glaucoma and/or normal tension glaucoma. These mutations will not be present in a sample obtained from a healthy person who is not at risk of onset of glaucoma. This polynucleotide sequence is the wild type human OPTN DNA that can be amplified to detect A619G/ G898A mutations indicative of the onset of glaucoma, given in an exemplification of the invention.
Sequence 1734 BP; 599 A; 347 C; 471 G; 317 T; 0 U; 0 Other;

Query Match	83.3%;	Score 1730.8;	DB 14;	Length 1734;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1732; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	311	ATGTCCTCATCAACTCTCAGCTGCTCTCACTGAAAGGAGGACAGCCCCAGTGAAGACACA	370	
Db	1	ATGTCCCATCAACTCTCAGCTGCTCTCACTGAAAGGAGGACAGCCCCAGTGAAGACACA	60	
Qy	371	GGAAATGGACCCCCCACTTGGCCCAACCCAAACCTGGACAGCTTTACCCCGGAGAGCTG	430	
Db	61	GGAAATGGACCCCCCACTTGGCCCAACCCAAACCTGGACAGCTTTACCCCGGAGAGCTG	120	
Qy	431	CTGCAGCAGATGAAGAGCTCTTGACCAAGNACCAACAGCTGAAAGAGCCATGAAGCTA	490	
Db	121	CTGCAGCAGATGAAGAGCTCTTGACCGAGAACCAACAGCTGAAAGAGCCATGAAGCTA	180	
Qy	491	AATAATCAAGCCATGAAGGGAGATTGTAGAGAGCTTTTCGGCTCGACAGAGAAACAGAA	550	
Db	181	AATAATCAAGCCATGAAGGGAGATTGTAGAGAGCTTTTCGGCTCGACAGAGAAACAGAA	240	
Qy	551	GAAGAACGCCAGTTTTTTTGAGATACAGAGCAAGAAGCAAAAGAGCGTCTTAATGGCCTTG	610	
Db	241	GAAGAACGCCAGTTTTTTTGAGATACAGAGCAAGAAGCAAAAGAGCGTCTTAATGGCCTTG	300	
Qy	611	AGTCATGAGAAATCAGAAATTTGAAGGAAGAGCTTGGAAAACTTAAAGGGAAATCAGAAAG	670	
Db	301	AGTCATGAGAAATCAGAAATTTGAAGGAAGAGCTTGGAAAACTTAAAGGGAAATCAGAAAG	360	
Qy	671	TCATCTGAGACCCCACTGATGACTCCAGGCTTCCAGGCGCGAGCGGAGCAGGAAAG	730	
Db	361	TCATCTGAGACCCCACTGATGACTCCAGGCTTCCAGGCGCGAGCGGAGCAGGAAAG	420	
Qy	731	GACCACTCAGGACCCAGGTGGTGAGGCTACAAGCAGAGAGGACAGCTGTTCGGGCATC	790	
Db	421	GACCACTCAGGACCCAGGTGGTGAGGCTACAAGCAGAGAGGACAGCTGTTCGGGCATC	480	
Qy	791	GTGCTCTGAACCTGACAGCTCAAGCTGAACCTCAGCGGCTCTCAGAAAGATTCTCTTTGTGAA	850	
Db	481	GTGCTCTGAACCTGACAGCTCAAGCTGAACCTCAGCGGCTCTCAGAAAGATTCTCTTTGTGAA	540	
Qy	851	ATTAGAGTGCTGAAGGAGAGCAAGAGGTCTAGTAAAGAAATCAAGACATAGTCTCTGGG	910	
Db	541	ATTAGAGTGCTGAAGGAGAGCAAGAGGTCTAGTAAAGAAATCAAGACATAGTCTCTGGG	600	

CC fertility, diabetes, adult respiratory distress syndrome, viral,
CC bacterial and parasitic infections. The nucleic acid sequences may be
CC used in chromosome mapping, identifying individual from minute biological
CC samples (tissue typing), and in forensic identification of a biological
CC sample. This sequence encodes a novel human protein (NOV)
XX
SQ Sequence 1908 BP; 637 A; 376 C; 505 G; 390 T; 0 U; 0 Other;

Query Match 81.3%; Score 1688.8; DB 8; Length 1908;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1693; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 378 GACCCGCCACCTGCCACCCCAACCTGGACAGCTTTACCCCGAGAGAGCTGCTGCAGC 437
DB 50 GATCCTCCACCTGGATCTCCCAACCTGGACAGCTTTACCCCGAGAGAGCTGCTGCAGC 109
QY 438 AGATGAAGAGCTCTGACCAAGAACCAACAGCTGAAAGAGCCATGAAGCTTAATAATC 497
DB 110 AGATGAAGAGCTCTGACCAAGAACCAACAGCTGAAAGAGCCATGAAGCTTAATAATC 169
QY 498 AAGCCATGAAGAGGAGATTTGAGGAGCTTTCCGCTGGACAGAGAAACAGAGGAAGAAC 557
DB 170 AAGCCATGAAGAGGAGATTTGAGGAGCTTTCCGCTGGACAGAGAAACAGAGGAAGAAC 229
QY 558 GCCAGTTTTTTGAGATCAGAGCAAGAGCAAGCAAGAGCGCTTAATGGCCCTTGAGTCATG 617
DB 230 GCCAGTTTTTTGAGATCAGAGCAAGAGCAAGCAAGAGCGCTTAATGGCCCTTGAGTCATG 289
QY 618 AGATGAGAAATTCAGAGAGAGCTTGGAAACTTAAAGGGAATCAGAAAGGTCATCTG 677
DB 290 AGATGAGAAATTCAGAGAGAGCTTGGAAACTTAAAGGGAATCAGAAAGGTCATCTG 349
QY 678 AGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAGCGAGCAGGAAAGGACCAAGC 737
DB 350 AGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAGCGAGCAGGAAAGGACCAAGC 409
QY 738 TCAGGACCCAGGTGGTGAAGCTCAAGCAGAGAGGAGCAGCTTTGGGCAATCGTGTCTG 797
DB 410 TCAGGACCCAGGTGGTGAAGCTCAAGCAGAGAGGAGCAGCTTTGGGCAATCGTGTCTG 469
QY 798 AACTGCACTCAAGCTCAAGCTCAAGCGGCTCCTCAGAGATTCCTTTGTTGAATTAGGA 857
DB 470 AACTGCACTCAAGCTCAAGCTCAAGCGGCTCCTCAGAGATTCCTTTGTTGAATTAGGA 529
QY 858 TGGCTGAAGGAGAGCAGAGAGGTCAGTAAAGAAATCAAGCATAGTCTCTGGGCCACGA 917
DB 530 TGGCTGAAGGAGAGCAGAGAGGTCAGTAAAGAAATCAAGCATAGTCTCTGGGCCACGA 589
QY 918 GAACAGTCTCCACTGGCAAGGCTTGTCTAAATATAGGAGCAGATCTGCAGATGGGGCCA 977
DB 590 GAACAGTCTCCACTGGCAAGGCTTGTCTAAATATAGGAGCAGATCTGCAGATGGGGCCA 649
QY 978 AGAATTAATTCAGACATGAGGAGTTACTGTAGCCAGCTCTGCTGTGCTTAAGGGAAG 1037
DB 650 AGAATTAATTCAGACATGAGGAGTTACTGTAGCCAGCTCTGCTGTGCTTAAGGGAAG 709
QY 1038 GGAATCAGAAGTGGAGAGCTTCAAGTTCAGCTCAAGAGGCGCCAAAGAGAGCTTTTCAG 1097
DB 710 GGAATCAGAAGTGGAGAGCTTCAAGTTCAGCTCAAGAGGCGCCAAAGAGAGCTTTTCAG 769
QY 1098 ATTTTGAAGAAAGAAACAAAGTAATCGTTCTGAGATTTGAACCCAGACAGAGGGGAGCAG 1157
DB 770 ATTTTGAAGAAAGAAACAAAGTAATCGTTCTGAGATTTGAACCCAGACAGAGGGGAGCAG 829
QY 1158 AGAAGAGATGATGAAGAGAAAGCCCGAGAGCTGTTGGAAGCGAAGTGAAGCACTGA 1217
DB 830 AGAAGAGATGATGAAGAGAAAGCCCGAGAGCTGTTGGAAGCGAAGTGAAGCACTGA 889
QY 1218 ACCTCCAGGTGACATCTCTGTTTAAGAGCTTCAAGAGGCTTCAACAAAACCTCAGCGAAG 1277
DB 890 ACCTCCAGGTGACATCTCTGTTTAAGAGCTTCAAGAGGCTTCAACAAAACCTCAGCGAAG 949
QY 1278 CTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCTAGGCGCTTTGAAGGAAATTTCTG 1337

DB 950 CTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAGGCGCTTTGAAAGGAAATTTCTG 1009
QY 1338 CAATTCCTCAGAGTTCGAATGAAGCAAGAGCTTTGTTTACTTAAACAAAAGTTAGAGC 1397
DB 1010 CAATTCCTCAGAGTTCGAATGAAGCAAGAGCTTTGTTTACTTAAACAAAAGTTAGAGC 1069
QY 1398 TACAAGTGAAGAGCATCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAGGATGAAA 1457
DB 1070 TACAAGTGAAGAGCATCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAGGATGAAA 1129
QY 1458 AGTCCAAATTTAACTGTGCTACAGATGACACAAACAGCTTTCTTCAAGAAACATTAATG 1517
DB 1130 AGTCCAAATTTAACTGTGCTACAGATGACACAAACAGCTTTCTTCAAGAAACATTAATG 1189
QY 1518 CATTGAAAAACAATTTGAGGAACTAACAGAAAGAGTCAGAAAAAGTGGACAGGCGAGTGC 1577
DB 1190 CATTGAAAAACAATTTGAGGAACTAACAGAAAGAGTCAGAAAAAGTGGACAGGCGAGTGC 1249
QY 1578 TGAAGGAACCTGAGTGAAGAACTGGAACTGGCAGAGAAAGGCTCTGGCTTCCAAAACAGCTGC 1637
DB 1250 TGAAGGAACCTGAGTGAAGAACTGGAACTGGCAGAGAAAGGCTCTGGCTTCCAAAACAGCTGC 1309
QY 1638 AATFGATGAATGAAGCAAAACCATTCGCCAAGCAGGAAGAGGACCTCGAAAACCATGACCA 1697
DB 1310 AATFGATGAATGAAGCAAAACCATTCGCCAAGCAGGAAGAGGACCTCGAAAACCATGACCA 1369
QY 1698 TCCTCAGGCTCAGATGAGAAAGTTTACTGTTCTGATTTTCTGCTGAAAGAGCAGCGAGAG 1757
DB 1370 TCCTCAGGCTCAGATGAGAAAGTTTACTGTTCTGATTTTCTGCTGAAAGAGCAGCGAGAG 1429
QY 1758 AGAAATTTCAATGAGGAAAGAGCAACTGGCATTGTCAGCTGGCAGTTCTGCTGAAAGAGA 1817
DB 1430 AGAAATTTCAATGAGGAAAGAGCAACTGGCATTGTCAGCTGGCAGTTCTGCTGAAAGAGA 1489
QY 1818 ATGATGCTTTGGAAGACGAGGAGGAGGAGCTCTTTGATGAGAGATGACAGAGTCTCATGGGG 1877
DB 1490 ATGATGCTTTGGAAGACGAGGAGGAGGAGCTCTTTGATGAGAGATGACAGAGTCTCATGGGG 1549
QY 1878 CGAAGCAAGTACTCTGACCCAGCAGGCTTACCTTTGTTCAAGAGGAGCTTGAGCAGCAGGG 1937
DB 1550 CGAAGCAAGTACTCTGACCCAGCAGGCTTACCTTTGTTCAAGAGGAGCTTGAGCAGCAGGG 1609
QY 1938 ACTGGGCGCAACAGCGGAATATTCGATTCATTCCTGCCCAAGTGTGGAGAGGTTCTGC 1997
DB 1610 ACTGGGCGCAACAGCGGAATATTCGATTCATTCCTGCCCAAGTGTGGAGAGGTTCTGC 1669
QY 1998 CTGACATAGACAGCTTACAGATTCAGCTGATGAGATGTCATCATTTAAGTGTGATGTATC 2057
DB 1670 CTGACATAGACAGCTTACAGATTCAGCTGATGAGATGTCATCATTTAAGTGTGATGTATC 1729
QY 2058 ACCTCCCAACAACTGTTGGT 2077
DB 1730 ACCTCCCAACAACTGTTGGT 1749

RESULT 20

ACD19488

ID ACD19488 standard; cdna; 1908 BP.

XX ACD19488;

XX AC

XX 25-AUG-2003 (first entry)

XX cDNA encoding novel human protein #168.

DE Human; NOV; gene therapy; endocrine related disease; diabetes;

XX metabolism-related disease; obesity; central nervous system disorder;

KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;

KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;

KW psoriasis; allergy; lupus erythematosus; asthma; cancer;

KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;

KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;

Db 1130 AGTCAAAATTAACGTGTGTCAGATGACACACAAAGCTTCTTCAAGAAATATAATG 1189
Qy 1518 CATTGAAAAAATTTGAGGAACTTAACAAGAAAAAGAGTCAGAAAAAGTGGACAGGGCAGTGC 1577
Db 1190 CATTGAAAAAATTTGAGGAACTTAACAAGAAAAAGAGTCAGAAAAAGTGGACAGGGCAGTGC 1249
Qy 1578 TGAAGGAATGAGTGAAGAACTGGAATGCGGAGAGAGGCTCTGGCTTCCAAAACAGTGC 1637
Db 1250 TGAAGGAATGAGTGAAGAACTGGAATGCGGAGAGAGGCTCTGGCTTCCAAAACAGTGC 1309
Qy 1638 AAATGGATGAATGAAGCAACCAATTCGCAAGCAGGAGAGGACTGGAAACCATGACCA 1697
Db 1310 AAATGGATGAATGAAGCAACCAATTCGCAAGCAGGAGAGGACTGGAAACCATGACCA 1369
Qy 1698 TCCTCAGGGCTCAGATGGAAGTTTACTGTCTTCTGATTTTTCATGCTGAAAGAGCAGCGAG 1757
Db 1370 TCCTCAGGGCTCAGATGGAAGTTTACTGTCTTCTGATTTTTCATGCTGAAAGAGCAGCGAG 1429
Qy 1758 AGAAATTCATGAGGAAAGAGCAACTGGCAATTCGAGCTGGAGTTCTGCTGAAAGAGA 1817
Db 1430 AGAAATTCATGAGGAAAGAGCAACTGGCAATTCGAGCTGGAGTTCTGCTGAAAGAGA 1489
Qy 1818 ATGATGCTTTCGAAGACGGAGCAGGAGTCTTGTGATGGAGATCAGAGTCTCATGGG 1877
Db 1490 ATGATGCTTTCGAAGACGGAGCAGGAGTCTTGTGATGGAGATCAGAGTCTCATGGG 1549
Qy 1878 CGAAGCAAGTGAATCTGACAGCAGGCTTACCTTGTTCGAAAGAGGAGCTGAGCAGGG 1937
Db 1550 CGAAGCAAGTGAATCTGACAGCAGGCTTACCTTGTTCGAAAGAGGAGCTGAGCAGGG 1609
Qy 1938 ACTGGCGCAACAGCGGAATATCCGATTCATCTGCCCCCAAGTGGAGAGGTTCTGC 1997
Db 1610 ACTGGCGCAACAGCGGAATATCCGATTCATCTGCCCCCAAGTGGAGAGGTTCTGC 1669
Qy 1998 CTGACATAGACACTTACAGATTCACGTGATGGATTCATTAAGTGTGATGATC 2057
Db 1670 CTGACATAGACACTTACAGATTCACGTGATGGATTCATTAAGTGTGATGATC 1729
Qy 2058 ACCTCCCCCAAACTGTTGGT 2077
Db 1730 ACCTCCCCCAAACTGTTGGT 1749
RESULT 21
ACD19486
ID ACD19486 standard; cDNA; 1858 BP.
XX AC ACD19486;
XX
XX 25-AUG-2003 (first entry)
XX cDNA encoding novel human protein #166.
XX Human; NOV; gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW stroke; infection; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2003023002-A2.
XX
XX 20-MAR-2003.
XX
XX 09-SEP-2002; 2002WO-US028539.
XX
XX

07-SEP-2001; 2001US-0318120P.
10-SEP-2001; 2001US-0318130P.
17-SEP-2001; 2001US-0318430P.
17-SEP-2001; 2001US-0322636P.
17-SEP-2001; 2001US-0322781P.
17-SEP-2001; 2001US-0322816P.
17-SEP-2001; 2001US-0322817P.
19-SEP-2001; 2001US-0323519P.
20-SEP-2001; 2001US-0323631P.
20-SEP-2001; 2001US-0323636P.
25-SEP-2001; 2001US-0324969P.
25-SEP-2001; 2001US-0325091P.
26-SEP-2001; 2001US-0324990P.
17-APR-2002; 2002US-0373212P.
06-SEP-2002; 2002US-00236177.

(CURA-) CURAGEN CORP.

Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
Lepley DM, Edinger SR, Burgess CE;

WPI; 2003-313242/30.

P-PSDB; ABO14793.

New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
and polynucleotides, useful in gene therapy, e.g. for treating or
preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
stroke or infections.

Claim 20; Page 434; 586pp; English.

The invention describes a new isolated polypeptide (NOVX). The NOVX
polypeptide, nucleic acid and antibody are useful as therapeutics,
particularly in the manufacture of a medicament for treating a syndrome
associated with a human disease, which includes a pathology associated
with NOVX polypeptide. The DNA encoding the protein is useful in gene
therapy for treating the disease or condition. In particular, the NOVX
polypeptide or polynucleotide is useful for treating endocrine/
metabolism-related diseases (e.g. obesity or diabetes), central nervous
system disorders (e.g. Alzheimer's disease, Parkinson's disease,
epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
asthma, inflammatory bowel disease, rheumatoid arthritis or
osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
prostate or brain cancers, or melanoma), liver diseases (e.g. liver
cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
These are also useful in developing powerful assay system for functional
analysis of various human disorders, as well as in diagnostic
applications, and for monitoring the effects of drugs during clinical
trials. This sequence encodes a novel human NOV protein

Sequence 1858 BP; 624 A; 360 C; 497 G; 377 T; 0 U; 0 Other;

Query Match 81.3%; Score 1687.8; DB 10; Length 1858;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1692; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 379 ACCCCCCCACTGGCCCCCACCACCACTGGACACGTTTACCCCCGAGGAGCTGTGCAGCA 438
Db 1 ATCTCTCCACTGGATCTCCCAAACTGGACAGTTTACCCCCGAGGAGCTGTGCAGCA 60

Qy 439 GATGAAAGAGCTCTTGACCAAGAACCAACCAAGCTGAAAGAGCCATGAAGCTAAATATCA 498
Db 61 GATGAAAGAGCTCTTGACCGAGAACCAACCAAGCTGAAAGAGCCATGAAGCTAAATATCA 120

Qy 499 AGCCATGAAAGGAGGAGATTGAGGAGCTTTCGGCTTGGACAGAGAAACAGAGGAGAACG 558
Db 121 AGCCATGAAAGGAGGAGATTGAGGAGCTTTCGGCTTGGACAGAGAAACAGAGGAGAACG 180

QY	559	CCAGTTTTTTTGATACAGAGCAAGAAAGACGCTCTAATGGCCTTGAGTCATGA	618	Db	1261	AATGATGAATGAAGCAAAACCATTTGCCAAGCAGGAAGGACCTTGGAAACCATGACCAT	1320
Db	181	CCAGTTTTTTTGATACAGAGCAAGAAAGACGCTCTAATGGCCTTGAGTCATGA	240	QY	1699	CCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGACGACGAGAGA	1758
QY	619	GAATGAGAAATTGAAGCAAGAGCTTGGAAATCTAAAGGGGAATCAGAAAGTCTCTCA	678	Db	1321	CCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGACGACGAGAGA	1380
Db	241	GAATGAGAAATTGAAGCAAGAGCTTGGAAATCTAAAGGGGAATCAGAAAGTCTCTCA	300	QY	1759	GAAATTTCAATGAGGAAAAAGGACAACTGGCATTGCACTGGGCAAGTTCTGCTGAAAGAGAA	1818
QY	679	GGACCCCACTGATGACTCCAGGCTTCCAGGGCCGGAAGCGGACAGCAAAAGACCAAGCT	738	Db	1381	GAAATTTCAATGAGGAAAAAGGACAACTGGCATTGCACTGGGCAAGTTCTGCTGAAAGAGAA	1440
Db	301	GGACCCCACTGATGACTCCAGGCTTCCAGGGCCGGAAGCGGACAGCAAAAGGACCAAGCT	360	QY	1819	TCATGCTTTGGAAGCGGAGGACGAGCTCTTGTGATGGAGATGCACAGTCTCATCGGGC	1878
QY	739	CAGGACCCAGGTGGTGAAGGCTA CAAGCAGAGAAAGGACAGCACTGTGTTGGGCATCGTGTGA	798	Db	1441	TCATGCTTTGGAAGCGGAGGACGAGCTCTTGTGATGGAGATGCACAGTCTCATCGGGC	1500
Db	361	CAGGACCCAGGTGGTGAAGGCTA CAAGCAGAGAAAGGACAGCACTGTGTTGGGCATCGTGTGA	420	QY	1879	GAGAACAACTGACTCTGACACGAGCGCTTACCTTGTTCAAAAGAGGAGCTGAGACACAGGA	1938
QY	799	ACTGCAAGCTCAAGCTGAACTTCAGCGGCTCTCAGAAAGATTCTTTGTTGAAATTAGGAT	858	Db	1501	GAGAACAACTGACTCTGACACGAGCGCTTACCTTGTTCAAAAGAGGAGCTGAGACACAGGA	1560
Db	421	ACTGCAAGCTCAAGCTGAACTTCAGCGGCTCTCAGAAAGATTCTTTGTTGAAATTAGGAT	480	QY	1939	CTGGCGGCAACAGCGGAATATTCCGATTCACTTCCTGCCCAAGTGTGGAGAGTTCTGCC	1998
QY	859	GGCTGAAGGAGACGAGAGGCTCAGTAAAGAAATCAAGCATAGTCCCTGGGCCACAGAG	918	Db	1561	CTGGCGGCAACAGCGGAATATTCCGATTCACTTCCTGCCCAAGTGTGGAGAGTTCTGCC	1620
Db	481	GGCTGAAGGAGACGAGAGGCTCAGTAAAGAAATCAAGCATAGTCCCTGGGCCACAGAG	540	QY	1999	TGACATAGACAGCTTACAGATTCACTGATGATTCATCATTTAAGTGTGATGATATCA	2058
QY	919	AACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGAGCAGATCTGCAATGGGCCCCAA	978	Db	1621	TGACATAGACAGCTTACAGATTCACTGATGATTCATCATTTAAGTGTGATGATATCA	1680
Db	541	AACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGAGCAGATCTGCAATGGGCCCCAA	600	QY	2059	CCTCCCCAAAACCTGTTGGT	2077
QY	979	GAATTAATTGCAATGAGAGATTAACTGTGAGCGCTCTCTGCTGCTTAAGGAAAGG	1038	Db	1681	CCTCCCCAAAACCTGTTGGT	1699
Db	601	GAATTAATTGCAATGAGAGATTAACTGTGAGCGCTCTCTGCTGCTTAAGGAAAGG	660	RESULT 22			
QY	1039	GAATCAAGAGTGGAGAGACTTCAAGTTGCACTCAAGGAGGCCAAAGAAAGAGTTTCAGA	1098	ID	ABX97628		
Db	661	GAATCAAGAGTGGAGAGACTTCAAGTTGCACTCAAGGAGGCCAAAGAAAGAGTTTCAGA	720	XX	ABX97628		
QY	1099	TTTTTGAAGAAACCAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAGCAGAGA	1158	AC	ABX97628;		
Db	721	TTTTTGAAGAAACCAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAGCAGAGA	780	DT	16-MAY-2003 (first entry)		
QY	1159	GAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGAAGCACTGAA	1218	XX	cDNA encoding novel human protein NOV30a.		
Db	781	GAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGAAGCACTGAA	840	DE	Human; NOV; adrenoleukodystrophy; congenital adrenal hyperplasia;		
QY	1219	CCTCCAGTGCATCTCTGTTTAAGGAGCTTCAAGAGCTCATAACAAAACCTCAGCGAAGC	1278	XX	haemophilia; hypercoagulation; autoimmune disease; allergy;		
Db	841	CCTCCAGTGCATCTCTGTTTAAGGAGCTTCAAGAGCTCATAACAAAACCTCAGCGAAGC	900	KW	immunodeficiency; transplantation; Von Hippel-Lindau syndrome;		
QY	1279	TGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTGAGGCCCTTGAAGGAAAAATTCGC	1338	KW	Alzheimer's disease; stroke; tuberosus sclerosis; hypercalcaemia;		
Db	901	TGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTGAGGCCCTTGAAGGAAAAATTCGC	960	KW	Parkinson's disease; Huntington's disease; cancer; fertility; diabetes;		
QY	1339	AATTCATCAGAGTTGAATGAAAGCAAGCTGTTGTTTACTTAAACAAAAGTTAGAGCT	1398	KW	adult respiratory distress syndrome; infection; tissue typing;		
Db	961	AATTCATCAGAGTTGAATGAAAGCAAGCTGTTGTTTACTTAAACAAAAGTTAGAGCT	1020	XX	forensic identification; gene; ss.		
QY	1399	ACAAGTGGAAAGCATGTATCAGAAATCAAAATGGACAGCTTAAACACAGAGATGAAA	1458	OS	Homo sapiens.		
Db	1021	ACAAGTGGAAAGCATGTATCAGAAATCAAAATGGACAGCTTAAACACAGAGATGAAA	1080	XX	WO200290500-A2.		
QY	1459	GTCCAAATTAATCTGTCTACAGATGACACACAAACAGCTTCTTCAAGAACATATATATGC	1518	PN	14-NOV-2002.		
Db	1081	GTCCAAATTAATCTGTCTACAGATGACACACAAACAGCTTCTTCAAGAACATATATATGC	1140	XX	02-MAY-2002; 2002WO-US014256.		
QY	1519	ATTGAAAAATTTGAGGAATTAACAAGAAAGAGTCAAGAAAAGTGACAGGGCAGTCT	1578	XX	03-MAY-2001; 2001US-0288395P.		
Db	1141	ATTGAAAAATTTGAGGAATTAACAAGAAAGAGTCAAGAAAAGTGACAGGGCAGTCT	1200	PR	07-MAY-2001; 2001US-0289087P.		
QY	1579	GAAGGAATCTGATCAAAAACCTGGAACTGGCAGAGAGCTCTGGCTTCCAAAACAGCTGCA	1638	PR	08-MAY-2001; 2001US-0289619P.		
Db	1201	GAAGGAATCTGATCAAAAACCTGGAACTGGCAGAGAGAGCTCTGGCTTCCAAAACAGCTGCA	1260	PR	09-MAY-2001; 2001US-0289817P.		
QY	1639	AATGGATGAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGACCTGGAAAACCATGACCAT	1698	PR	11-MAY-2001; 2001US-0290194P.		
				PR	14-MAY-2001; 2001US-0290753P.		
				PR	15-MAY-2001; 2001US-0291189P.		
				PR	21-MAY-2001; 2001US-0292374P.		
				PR	23-MAY-2001; 2001US-0293107P.		
				PR	25-MAY-2001; 2001US-0293747P.		
				PR	29-MAY-2001; 2001US-0294110P.		
				PR	30-MAY-2001; 2001US-0294434P.		
				PR	10-SEP-2001; 2001US-0318346P.		
				PR	17-SEP-2001; 2001US-0322646P.		

PR 01-MAY-2002; 2002US-00136728.
XX (CURA-) CURAGEN CORP.
XX Spvtek KA, Li L, Edinger SR, Stone DJ, Guo X, Anderson DW;
PI Patturajan M, Gerlach VL, Taupier RJ, Pena CEA, Padigaru M;
PI Kekuda R, Gorman L, Zernhusen BD, Smithson G, Macdougall JR;
PI Mezes PS, Peyman JA, Zhong M;
XX WPI; 2003-103511/09.
DR P-PSDB; ABU65253.
XX
XX New NOVX polypeptides and polynucleotides useful for treating or
PT preventing e.g. congenital adrenal hyperplasia, hemophilia,
PT hypercoagulation, autoimmune disease, allergies, immunodeficiencies,
PT transplantation.
XX
XX Claim 1; Page 162; 300pp; English.
XX
XX The invention describes an isolated polypeptide, NOVX, comprising a
CC sequence or a mature form of one of 21 51-1543 residue amino acid
CC sequences (p1-p21), given in the specification. The NOVX polypeptides,
CC polynucleotides and antibodies are useful in the manufacture of a
CC medicament for treating or preventing e.g. adrenoleukodystrophy,
CC congenital adrenal hyperplasia, haemophilia, hypercoagulation, autoimmune
CC disease, allergies, immunodeficiencies, transplantation, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cancer,
CC fertility, diabetes, adult respiratory distress syndrome, viral,
CC bacterial, and parasitic infections. The nucleic acid sequences may be
CC used in chromosome mapping, identifying individual from minute biological
CC samples (tissue typing), and in forensic identification of a biological
CC sample. This sequence encodes a novel human protein (NOV)
XX
SQ Sequence 1908 BP; 633 A; 380 C; 505 G; 390 T; 0 U; 0 Other;

Query Match 79.3%; Score 1646.6; DB 8; Length 1908;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 7; Indels 18; Gaps 1;

QY 379 ACCCCCCACCTGGCCACCACCAACCTGGACACGTTTACCCCGGAGAGCTGTCGACGA 438
DB 69 ATCCTCCACCTGGATCTCCCAACCTGGACACATTTACCCCGGAGAGCTGTCGACGA 128

QY 439 GATGAAGAGCTCTCTGACCAAGAACACACAGCTGAAAGAGCCATGAAGCTAAATATCA 498
DB 129 GATGAAGAGCTCTCTGACCAAGAACACACAGCTGAAAGAGCCATGAAGCTAAATATCA 188

QY 499 AGCCATGAAGAGGAGATTGAGGAGCTTTCGGCTCGACAGAGAAACAGAGAAAGAACG 558
DB 189 AGCCATGAAGAGGAGATTGAGGAGCTTTCGGCTCGACAGAGAAACAGAGAAAGAACG 248

QY 559 CAGTTTTTTGAGATACAGAGCAAGAAAGCAAGAGCGTCTAATGSCCTTGAGTCATGA 618
DB 249 CAGTTTTTTGAGATACAGAGCAAGAAAGCAAGAGCGTCTAATGSCCTTGAGTCATGA 308

QY 619 GAATGAGAAATTCGAGAGAGCTTGAAACCTGAAAGGAAATCGAAAGGTCATCTGA 678
DB 309 GAATGAGAAATTCGAGAGAGCTTGAAACCTGAAAGGAAATTCGAAAGGTCATCTGA 368

QY 679 GGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAGAGCGAGAGAAAGAACCAAGCT 738
DB 369 GGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAGAGCGAGAGAAAGAACCAAGCT 428

QY 739 CAGGACCCAGGTGTAGGCTTAAAGAGAGAGAGAGAGCTGTTGGGCAATCGTGTCTGA 798
DB 429 CAGGACCCAGGTGTAGGCTTAAAGAGAGAGAGAGAGCTGTTGGGCAATCGTGTCTGA 488

QY 799 ACTGCACTCAAGCTGAAGCTCAGCGGCTCCTCAGAGATTCTTTGTTGAAATAGGAT 858
DB 489 ACTGCACTCAAGCTGAAGCTCAGCGGCTCCTCAGAGATTCTTTGTTGAAATAGGAT 548

QY 859 GGCTGAAGGAGAGAGAGAGGGTTCAGTAAAGAAATCAAGCATAGTCTCTGGGCCACGAG 918

DB 549 GGCTGAAGGAGAGAGAGGGTCAGTAAAGAAATCAAGCATAGTCTCTGGGCCACGAG 608
QY 919 AACAGTCTCCACTGGCACGGCATTTGCTAAATATATAGAGAGAGATCTCCAGATGGGGCAA 978
DB 609 AACAGTCTCCACTGGCAC-----GAGCAGATCTTCAGATGGGGCAA 650
QY 979 GAATTAATCTCGAAACATGAGGAGTTAACTGTGAGCCAGCTCTCTGCTGCTCCTAAGGGAAG 1038
DB 651 GAATTAATCTCGAAACATGAGGAGTTAACTGTGAGCCAGCTCTCTGCTGCTCCTAAGGGAAG 710
QY 1039 GAATCAGAGGTGGAGAGACTTGAAGTTGCATCAAGAGGCGCCAAAGAAAGAGTTTCAGA 1098
DB 711 GAATCAGAGGTGGAGAGACTTGAAGTTGCATCAAGAGGCGCCAAAGAAAGAGTTTCAGA 770
QY 1099 TTTTGAAGAAAGAAACAAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGACACAGA 1158
DB 771 TTTTGAAGAAAGAAACAAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGACACAGA 830
QY 1159 GAAAGAGAAATGATCAAGAGAGAAAGGCCCGGAGACTGTTGGAAGCGAACTGGAAGCACTGAA 1218
DB 831 GAAAGAGAAATGATCAAGAGAGAAAGGCCCGGAGACTGTTGGAAGCGAACTGGAAGCACTGAA 890
QY 1219 CTTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATACAAAACTCAGCGAAGC 1278
DB 891 CTTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATACAAAACTCAGCGAAGC 950
QY 1279 TGAGCTTAATCAAGAGAGACTTCAAGAAAGTCTCAGGCCCTTGAAGGAAATTTCTGC 1338
DB 951 TGAGCTTAATCAAGAGAGACTTCAAGAAAGTCTCAGGCCCTTGAAGGAAATTTCTGC 1010
QY 1339 AATTCATCAGAGCTGAATGAAAGCAAGAGCTTGTATTATCTAAACAAAAGTTAGAGCT 1398
DB 1011 AATTCATCAGAGCTGAATGAAAGCAAGAGCTTGTATTATCTAAACAAAAGTTAGAGCT 1070
QY 1399 ACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAGATGAAAA 1458
DB 1071 ACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAGATGAAAA 1130
QY 1459 GTCCAAATTAATCTGCTACAGATGACACACACAAAGCTTCTTCAAGAAATATATATGC 1518
DB 1131 GTCCAAATTAATCTGCTACAGATGACACACAAAGCTTCTTCAAGAAATATATATGC 1190
QY 1519 ATTGAAAAATTAATGAGGACTTAAAGAAAGAGTCAAGAAAAAGTGACAGGCGAGTGTCT 1578
DB 1191 ATTGAAAAATTAATGAGGACTTAAAGAAAGAGTCAAGAAAAAGTGACAGGCGAGTGTCT 1250
QY 1579 GAAGGAATCTGAGTGAAGAAATCTGGAACCTGGCAGAGAAAGGCTCTGGCTTCCAAAACAGCTGCA 1638
DB 1251 GAAGGAATCTGAGTGAAGAAATCTGGAACCTGGCAGAGAAAGGCTCTGGCTTCCAAAACAGCTGCA 1310
QY 1639 AATGGATGAAATGAAGCAAAACCAATTCGCAAGAGAGAGGAGCTGCGAAACCAATGACCAT 1698
DB 1311 AATGGATGAAATGAAGCAAAACCAATTCGCAAGAGAGAGGAGCTGCGAAACCAATGACCAT 1370
QY 1699 CCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCTGCTGAAAGACAGCGAGAGA 1758
DB 1371 CCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCTGCTGAAAGACAGCGAGAGA 1430
QY 1759 GAAATTCATGAGAAAAAGGAGCACTGGCATTTGCAGCTGGCAGTTCGCTGAAAGAGAA 1818
DB 1431 GAAATTCATGAGAAAAAGGAGCACTGGCATTTGCAGCTGGCAGTTCGCTGAAAGAGAA 1490
QY 1819 TGATGCTTTCGAAGACGAGGCGAGGAGCTCTTTGATGGAGATGACAGATGCTCATGGGGC 1878
DB 1491 TGATGCTTTCGAAGACGAGGCGAGGAGCTCTTTGATGGAGATGACAGATGCTCATGGGGC 1550
QY 1879 GAGAAACAGTGTCTGACCGCAGGCTTACCTTGTTCAAAGAGGAGCTGAGACAGGGA 1938
DB 1551 GAGAAACAGTGTCTGACCGCAGGCTTACCTTGTTCAAAGAGGAGCTGAGACAGGGA 1610
QY 1939 CTGGCGGCAACAGCGGAATATTCCGATTTCCTGCCCCCAAGTGTGGAGAGTTCTTGGC 1998

Db 1611 CTGGCGCAACAGCGGAATATTCGGATTCTCTGCCCCCAAGTGTGGAGAGTTCTGCC 1670

Qy 1999 TGACATAGACACGTTACAGATTACACGTTACGTTGATTCATCATTTAGTGTGATGATCA 2058

Db 1671 TGACATAGACACGTTACAGATTACGTTGATTCATCATTTAGTGTGATGATCA 1730

Qy 2059 CTCTCCCAAACTGTTGGT 2077

Db 1731 CTCTCCCAAACTGTTGGT 1749

RESULT 23

ACD19485

ID ACD19485 standard; cDNA; 1908 BP.

XX ACD19485;

XX ACD19485;

DT 25-AUG-2003 (first entry)

XX cDNA encoding novel human protein #165.

XX Human; NOV; gene therapy; endocrine related disease; diabetes;

KW metabolism-related disease; obesity; central nervous system disorder;

KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;

KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;

KW psoriasis; allergy; lupus erythematosus; asthma; cancer;

KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;

KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;

KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;

KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;

KW stroke; infection; gene; ss.

XX Homo sapiens.

OS

XX

XX WO2003023002-A2.

XX

PD 20-MAR-2003.

XX

PF 09-SEP-2002; 2002WO-US028539.

XX

PR 07-SEP-2001; 2001US-0318120P.

PR 07-SEP-2001; 2001US-0318130P.

PR 10-SEP-2001; 2001US-0318430P.

PR 17-SEP-2001; 2001US-0322636P.

PR 17-SEP-2001; 2001US-0322781P.

PR 17-SEP-2001; 2001US-0322816P.

PR 17-SEP-2001; 2001US-0322817P.

PR 19-SEP-2001; 2001US-0323519P.

PR 20-SEP-2001; 2001US-0323631P.

PR 20-SEP-2001; 2001US-0323636P.

PR 25-SEP-2001; 2001US-0324969P.

PR 25-SEP-2001; 2001US-0325091P.

PR 26-SEP-2001; 2001US-0324990P.

PR 17-APR-2002; 2002US-0373212P.

PR 06-SEP-2002; 2002US-00236177.

XX (CURA-) CURAGEN CORP.

PA

XX

XX Spytak KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;

PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;

PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;

PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;

PI Lepley DM, Edinger SR, Burgess CE;

XX

XX WPI; 2003-313242/30.

DR

DR P-PSDB; ABO14792.

XX

XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)

PT and polynucleotides, useful in gene therapy, e.g. for treating or

PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,

PT stroke or infections.

XX

XX Claim 20; Page 433-434; 586pp; English.

PS

XX

CC The invention describes a new isolated polypeptide (NOVX). The NOVX

CC polypeptide, nucleic acid and antibody are useful as therapeutics,

CC particularly in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, which includes a pathology associated

CC with NOVX polypeptide. The DNA encoding the protein is useful in gene

CC therapy for treating the disease or condition. In particular, the NOVX

CC polypeptide or polynucleotide is useful for treating endocrine/

CC metabolism-related diseases (e.g. obesity or diabetes), central nervous

CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,

CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune

CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,

CC asthma, inflammatory bowel disease, rheumatoid arthritis or

CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,

CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver

CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),

CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).

CC These are also useful in developing powerful assay system for functional

CC analysis of various human disorders, as well as in diagnostic

CC applications, and for monitoring the effects of drugs during clinical

CC trials. This sequence encodes a novel human NOV protein

XX

SQ Sequence 1908 BP; 633 A; 380 C; 505 G; 390 T; 0 U; 0 Other;

Qy Query Match 79.3%; Score 1646.6; DB 10; Length 1908;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 1674; Conservative 0; Mismatches 7; Indels 18; Gaps 1;

Qy 379 ACCCCCCCACTGGCCGCCCAAACTCGACACGTTTACCCCGAGAGCTGCTGAGCA 438

Db 69 ATCTCTCCCACTGGGATCTCCCAAACTCGACACATTTACCCCGAGAGCTGCTGAGCA 128

Qy 439 GATGAAAGAGCTCTGACCAAGAACCCACCAGCTGAAAGAGCCATGAATTAATCA 498

Db 129 GATGAAAGAGCTCTGACCGAGAACCCACCAGCTGAAAGAGCCATGAATTAATCA 188

Qy 499 AGCCATGAAAGGAGATTGTGAGAGCTTTCCGCTGGACAGAGAAACAGAGGAAGAACG 558

Db 189 AGCCATGAAAGGAGATTGTGAGAGCTTTCCGCTGGACAGAGAAACAGAGGAAGAACG 248

Qy 559 CCAGTTTTTTTGATACACAGCAAGAACCAAGAGCGCTTAATGCGCTTGAGTCATGA 618

Db 249 CCAGTTTTTTTGATACACAGCAAGAACCAAGAGCGCTTAATGCGCTTGAGTCATGA 308

Qy 619 GAATCAGAAATTCAGAGAGAGCTTTGGAATACTAAAGGGAATTCAGAAAGTCACTGA 678

Db 309 GAATGAGAAATTCAGAGAGAGCTTTGGAATACTAAAGGGAATTCAGAAAGTCACTGA 368

Qy 679 GGACCCCACTGATGATCTCAGGCTTCCAGGGCCGAGCGGAGCAGAGAACCCAGCT 738

Db 369 GGACCCCACTGATGATCTCAGGCTTCCAGGGCCGAGCGGAGCAGAGAACCCAGCT 428

Qy 739 CAGGACCCAGGTGCTGAGGCTCAAGCAGAGAGGCGAGCCTCTTGGGCACTGCTGCTGA 798

Db 429 CAGGACCCAGGTGCTGAGGCTCAAGCAGAGAGGCGAGCCTCTTGGGCACTGCTGCTGA 488

Qy 799 ACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAGAGATTCTTTGTTGAAATTAGGAT 858

Db 489 ACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAGAGATTCTTTGTTGAAATTAGGAT 548

Qy 859 GCGTGAAGAGAGAGCAGAGAGGTGCTGATTAAGAAATCAAGCATAGTCTTGGGCCACGAG 918

Db 549 GCGTGAAGAGAGAGCAGAGAGGTGCTGATTAAGAAATCAAGCATAGTCTTGGGCCACGAG 608

Qy 919 AACAGTCTCCACTGGCAGCGCATTTGCTAAATATAGAGCAGATCTGCAGATGGGGCAA 978

Db 609 AACAGTCTCCACTGGGCAC-----GAGCAGATCTGCAGATGGGGCAA 650

Qy 979 GAATTACTTCGAACATGAGAGGTTAACTGTGAGCCAGCTCTCTGTCCTAAGGGAAGG 1038

Db 651 GAATTACTTCGAACATGAGAGGTTAACTGTGAGCCAGCTCTCTGTCCTAAGGGAAGG 710

Qy 1039 GAATCAGAGGTGAGAGACTTGAAGTTGCACCTCAAGGAGGCCCAAGAGAGGTTTTCAGA 1098

CC	asthma, inflammatory bowel disease, rheumatoid arthritis or osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian, prostate or brain cancers, or melanoma), liver diseases (e.g. liver cirrhosis), lung diseases (emphysema or obstructive pulmonary disease), haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic). These are also useful in developing powerful assay system for functional analysis of various human disorders, as well as in diagnostic CC applications, and for monitoring the effects of drugs during clinical CC trials. This sequence encodes a novel human NOV protein	
XX		
SQ	Sequence 1857 BP; 620 A; 363 C; 497 G; 377 T; 0 U; 0 Other;	
Query Match	79.0%; Score 1641.8; DB 10; Length 1857;	
Best Local Similarity	98.5%; Pred. No. 0;	
Matches 1674; Conservative	0; Mismatches 7; Indels 18; Gaps 1;	
QY	379 ACCCCGCCACCTGGCCCAACCAACCTGGACACGTTTACCCCGGAGAGCTGCTGCAGCA	438
DB	18 ATCTCCCACTGGATCTCCCAACCTGGACACATTTACCCCGGAGAGCTGCTGCAGCA	77
QY	439 GATGAAGAGCTCCTGACCAAGAACCAACAGCTGAAAGAGCCATGAAGCTAAATATCA	498
DB	78 GATGAAGAGCTCCTGACCGAGAACCAACAGCTGAAAGAGCCATGAAGCTAAATATCA	137
QY	499 AGCCATGAAGGAGATTTGAGAGCTTTGCGCCTGGACAGAGAAACAGAGGAAGACG	558
DB	138 AGCCATGAAGGAGATTTGAGAGCTTTGCGCCTGGACAGAGAAACAGAGGAAGAACG	197
QY	559 CCGATTTTGTAGTACAGAGCAAGAGCAAGAGAGCGCTTAATGCGCTTGAGTCAATGA	618
DB	198 CCGATTTTGTAGTACAGAGCAAGAGCAAGAGAGCGCTTAATGCGCTTGAGTCAATGA	257
QY	619 GAATGAGAAATCAAGAGAGAGCTTGGAACCTAAAGGGAAATCAGAAAGTCACTGA	678
DB	258 GAATGAGAAATCAAGAGAGAGCTTGGAACCTAAAGGGAAATCAGAAAGTCACTGA	317
QY	679 GGAACCCACATGATGACTCCAGGCTTCCAGGGCGGAGCGGAGAGGAAAGACCGCT	738
DB	318 GGAACCCACATGATGACTCCAGGCTTCCAGGGCGGAGCGGAGAGGAAAGACCGCT	377
QY	739 CAGGACCCAGTGTGAGGCTTACAAGCAGAGAGAGGACGCTGTGGGCATCGTCTGA	798
DB	378 CAGGACCCAGTGTGAGGCTTACAAGCAGAGAGAGGACGCTGTGGGCATCGTCTGA	437
QY	799 ACTGCAGCTCAAGCTGAACCTCCAGCGGCTCCTCAGAGATTCCTTTGTTGAAATTAGAT	858
DB	438 ACTGCAGCTCAAGCTGAACCTCCAGCGGCTCCTCAGAGATTCCTTTGTTGAAATTAGAT	497
QY	859 GGCTGAAGGAGAGCAGAGGGTCAGTAAAGAAATCAAGCATAGTCTGGGCCACGAG	918
DB	498 GGCTGAAGGAGAGCAGAGGGTCAGTAAAGAAATCAAGCATAGTCTGGGCCACGAG	557
QY	919 AACAGTCTCCACTGGCAGCGCATTTGCTAAATATAGGAGCAGATCTGCAGATGGGCCAA	978
DB	558 AACAGTCTCCACTGGCAC-----GAGCAGATCTGCAGATGGGCCAA	599
QY	979 GAATTACTTCGAACATCAGGAGTTAACTGTGAGCAGCTCCTGTGCTCCTAAAGGGAAGG	1038
DB	600 GAATTACTTCGAACATCAGGAGTTAACTGTGAGCAGCTCCTGTGCTCCTAAAGGGAAGG	659
QY	1039 GAATCAGAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAAGAAAGAGTTTCAGA	1098
DB	660 GAATCAGAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAAGAAAGAGTTTCAGA	719
QY	1099 TTTTGAAGAAAGAAACAGATTAATCGTTCTGAGATTGAAACCCAGACAGGGGAGCAGAGA	1158
DB	720 TTTTGAAGAAAGAAACAGATTAATCGTTCTGAGATTGAAACCCAGACAGGGGAGCAGAGA	779
QY	1159 GAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGGAAAGCACTGAA	1218
DB	780 GAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGGAAAGCACTGAA	839
QY	1219 CCTCCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAGCGAAGC	1278

DB	840	CCTCAGGTGACATCTCTGTTAAGAGGCTTCAAGAGGCTCATACAAAACCTCAGCGAAGC	899
QY	1279	TGAGCTTAATGAAGAAGAGACTTTCAAGAAAGTGTCAAGGCCCTTGAAGAGAAAATTTCTGC	1338
DB	900	TGAGCTTAATGAAGAAGAGACTTTCAAGAAAGTGTCAAGGCCCTTGAAGAGAAAATTTCTGC	959
QY	1339	AATTCCATCAGAGTTGAATGAAGAAAGCAGAGCTTGTGTTTATCTAAACAAAAGTTAGAGCT	1398
DB	960	AATTCCATCAGAGTTGAATGAAGAAAGCAGAGCTTGTGTTTATCTAAACAAAAGTTAGAGCT	1019
QY	1399	ACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACACAGAGATCAAAA	1458
DB	1020	ACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACACAGAGATCAAAA	1079
QY	1459	GTCCAAATTAACCTGTGCTACAGATGACACACAACAGCTTCTTCAAGAACATTAATATGC	1518
DB	1080	GTCCAAATTAACCTGTGCTACAGATGACACACAACAGCTTCTTCAAGAACATTAATATGC	1139
QY	1519	ATTGAAAAAATTTGAGGAACTTAACAAGAAAAAGTCAAGAAAAAGTGCACAGGCGAGTGT	1578
DB	1140	ATTGAAAAAATTTGAGGAACTTAACAAGAAAAAGTCAAGAAAAAGTGCACAGGCGAGTGT	1199
QY	1579	GAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCTGGCTTCCAAAACAGCTGCA	1638
DB	1200	GAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCTGGCTTCCAAAACAGCTGCA	1259
QY	1639	AATGATGAATTAAGCAACCAATTCGCAAGCAGGAGAGGACCTGGAAACCAATGACCAT	1698
DB	1260	AATGATGAATTAAGCAACCAATTCGCAAGCAGGAGAGGACCTGGAAACCAATGACCAT	1319
QY	1699	CCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCACTGCTGAAAGACGCGAGAGA	1758
DB	1320	CCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCACTGCTGAAAGACGCGAGAGA	1379
QY	1759	GAAATTCATGAGGAAAAAGAGCAACTGGCATTGCGAGCTGGCAGTCTGCTGAAAAGAGAA	1818
DB	1380	GAAATTCATGAGGAAAAAGAGCAACTGGCATTGCGAGCTGGCAGTCTGCTGAAAAGAGAA	1439
QY	1819	TGATGCTTTGGAAGACGAGGCGAGGAGTCTTTGATGGAGATCAGAGTCTCATGGGGC	1878
DB	1440	TGATGCTTTGGAAGACGAGGCGAGGAGTCTTTGATGGAGATCAGAGTCTCATGGGGC	1499
QY	1879	GAGAAACAGTGACTGACAGCAGGCTTACCTGTTTCAAGAGAGGCTCAGGACAGGGA	1938
DB	1500	GAGAAACAGTGACTGACAGCAGGCTTACCTGTTTCAAGAGAGGCTCAGGACAGGGA	1559
QY	1939	CTGGCGGCAACAGCGGAATATTCCGATTTCATTCTGCCCCAAGTGTGGAGAGGTTCTGCC	1998
DB	1560	CTGGCGGCAACAGCGGAATATTCCGATTTCATTCTGCCCCAAGTGTGGAGAGGTTCTGCC	1619
QY	1999	TGACATAGACACGTTTACAGATTCACGTTGATGGATTCGATCATTTAAGTGTGATGATCA	2058
DB	1620	TGACATAGACACGTTTACAGATTCACGTTGATGGATTCGATCATTTAAGTGTGATGATCA	1679
QY	2059	CCTCCCCAAAACCTGTTGGT 2077	
DB	1680	CCTCCCCAAAACCTGTTGGT 1698	
RESULT 25			
ADO43193			
ID	ADO43193	standard; cDNA; 1925 BP.	
XX	ADO43193;		
AC	ADO43193;		
XX	29-JUL-2004	(first entry)	
DT	29-JUL-2004	(first entry)	
XX	Macaque optineurin cDNA.		
DE	Mouse; optineurin; glaucoma; diagnosis; gene therapy; ophthalmological;		
XX	gene; ss.		
KW			
XX			

|||||
1438 AGGACCTGGAAACCATGACCGTCTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTC 1497
QY 1737 ATGCTGAAGAGCAGCAGAGAGAAATTCATGAGGAAAGGACAACTGGCATTGCAGC 1796
Db 1498 ATGCTGAAGAGCAGCAGAGAGAAATTCATGAGGAAAGGACAACTGGCATTGCAGC 1557
QY 1797 TGGCAGTTCTGCTGAAGAGAAATGATGCTTTTTCGAAGACGGAGCAGGCGATCCTTGATCG 1856
Db 1558 TGGCAGTTTCTGCTGAAGAGAAATGATGCTTTTTCGAAGATGGAGCAGGCGATCCTTGATCG 1617
QY 1857 AGATGACAGTCTGATGGGCGAGAACAGTACTCTGACACGAGGCTTACCTTGTTTC 1916
Db 1618 AGATGACAGTCTGATGGGCGAGAACAGTACTCTGACACGAGGCTTACCTTGTTTC 1677
QY 1917 AAGAGGAGCTGAGGACAGGAGCTGGCGCAACAGCGGAATATTCGATTTCATTCCTGCC 1976
Db 1678 AAGAGGAACTGAGGACAGGAGCTGGCGCAACAGCGGAATATTCGATTTCATTCCTGCC 1737
QY 1977 CCAAGTGTGAGAGGTTCTGCTTGACATAGACACGTTTACAGATTTCAGTGGATTGCA 2036
Db 1738 CCAAGTGTGAGAGGTTCTGCTTGACATAGACACGTTTACAGATTTCAGTGGATTGCA 1797
QY 2037 TCATTTAAGTGTGATGATATCACTTCCCAAAACTGTTGGT 2077
Db 1798 TCATTTAGTGTGATGATGATCACTTCCCAAAACTGTTGGT 1838

RESULT 26
AD043199
ID AD043199 standard; cdna; 1799 BP.
XX AC AD043199;
XX DT 29-JUL-2004 (first entry)
XX DE Pig optineurin cdna.
XX KW Pig; optineurin; glaucoma; diagnosis; gene therapy; ophthalmological;
XX KW gene; ss.
XX OS Sus scrofa.
XX FH Key Location/Qualifiers
XX FT 1. 1725
XX FT CDS /*tag= a
XX PN WO2004039312-A2.
XX PD 13-MAY-2004.
XX PF 26-JUN-2003; 2003WO-US020165.
XX PR 25-OCT-2002; 2002US-00281457.
XX PA (UYCO-) UNIV CONNECTICUT.
XX PA (SCEO-) ST GEORGES ENTERPRISES LTD.
XX PI Sarfarazi M, Rezaie T, Child AH;
XX WPI; 2004-376046/35.
XX P-PSDB; AD043200.
XX PT New optineurin nucleic acids and polypeptides having sequence
XX PT alterations, which indicates the presence of an optineurin-associated
XX PT (risk of) glaucoma, useful for detecting or diagnosing glaucoma, e.g.
XX PT primary open angle glaucoma.
XX PS Example; SEQ ID NO 17; 110pp; English.
XX CC The present sequence is that of pig optineurin (optic neuropathy inducing
XX CC protein) cdna. Comparison of protein sequences, and Western blotting
XX CC results, indicate conservation of optineurin as shown in human AD043184,

macaque, mouse, rat, pig and cattle sequences. Overall, human optineurin
CC has 78-85% sequence identity with its homologues in mouse, rat, pig and
CC cattle, and 96% identity with macaque. Mutation of the human optineurin
CC gene is associated with primary open angle glaucoma. 2 Mutations, E50K
CC and M98K, found in human optineurin are also found in macaque,
CC corroborating evidence that this mutation is a risk factor for glaucoma.
CC The E50K mutation is further conserved in mouse and bovine. Methods of
CC detection, prognosis and diagnosis of the presence or absence of
CC optineurin-associated glaucoma or of an optineurin-associated increased
CC risk of glaucoma are described, in which a sample is tested for the
CC presence of certain gene sequence alterations or for alterations in the
CC expression or activity of the optineurin protein. Also described are
CC methods of therapy of glaucoma, including gene therapy methods.
XX
SQ Sequence 1799 BP; 636 A; 366 C; 454 G; 343 T; 0 U; 0 Other;
Query Match 62.2%; Score 1291; DB 12; Length 1799;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1495; Conservative 0; Mismatches 245; Indels 27; Gaps 3;
QY 311 ATGTCCCATCAACCTCTCAGCTGCTCTACTGAAAGAGGAGCAGCCCTGTAAGACACA 370
Db 1 ATGTCCCATCAACCTCTGAGCTGCTCTACTGAGAGGGGACAGCCCCACCGAAACACA 60
QY 371 GGAATGGACCCCCCACCCTGCGCCACCCAAACCTTGACACGTTTACCCCGAGGAGCTG 430
Db 61 GGAATGGACCCCCCACCCTCTGGCTCACCCAAACCTTGACACGTTTACCCCGAGGAGCTG 120
QY 431 CTGACGACATGAAGAGCTCTCTGACCAAGAACCCACGCTGAAAGAACCATGAAGCTA 490
Db 121 CTGACGACATGAGAGGCTTCTATCTGAGAACCATCAGCTGAAAGAGCCATGAAGCTA 180
QY 491 AATAATCAAGCCATGAAGGGAGATTTGAGGAGCTTTGCGCCTTGGACAGAGAACAGAG 550
Db 181 AATAATCAAGCTATGAAGGGGATTTGAGGAGCTTTGAGGAGCTTTGAGGAGCTTTGAG 240
QY 551 GAAGAACGCGAGTTTTTTGAGATACAGAGCAAGAGCAAGAGAGGCTCTAATGGCGCTTG 610
Db 241 GAAGAACGCGCTTTTTTTGAGACCCAGAGCAAGAGAGGAGGAGGCTTAAAGGCTCTG 300
QY 611 AGTCATGAGAAATGAGAAATTTGAAGAGAGCTTGAAGAACTTAAAGGAAATCAGAAAG 670
Db 301 AGTCATGAGAAATGAGAAATTTGAAGCAAGAACTTGAAGAACTTAAAGGAAATCAGAAAG 360
QY 671 TCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCCGAGCGGAGCAGGAAAG 730
Db 361 TCATTTGAGGACCTCACTGGGGACCCCGAGGCTCCCAAGGCGGAGCAGAACAGGAGTA 420
QY 731 GACGAGCTCAGGACCCAGGCTGAGGCTACAAGCAGAGAGGAGCAGACCTGTTGGGCATC 790
Db 421 GAAACAGCTGAAGACCCAGGCTGAGGCTTCAAGCTGAAAGGCGGATCTGCTGGGCATC 480
QY 791 GTGCTCTGAAGCTGAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAAGATTCCTTTGTA 850
Db 481 GTGCTCTGAAGCTGAGCTCAAGCTGAACTCAGTGGGCTCTCTGAGACTCTTTGTTGAA 540
QY 851 ATTAGGATGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCAAGCATAGTCTCTGGG 910
Db 541 ATCAGGATGCTGAGGAGAGCAGATGACGAAATGAAGGAAATCAAGCAAGCTCTGGG 600
QY 911 CCCACGAGAACAGTCTCCACTGCGACGGCATTTCTTAAATATAGGAGCAGATCTGCAGAT 970
Db 601 CCCATGAAGAACTGATTCATTTGACAC-----GACAAATCTGCAGAA 642
QY 971 GGGGCCAAGAAATTAATTCGAAACATGAGGAGTTAACTGTGAGCCAGCTCTCTGCTGCTA 1030
Db 643 GGTACCAAGAAATTAATTTGAAATTTGAGAAATTAATGTGAGCCAGCTCTCTGCTGCTA 702
QY 1031 AGGGAAGGGAATCAGAGAGGTTGAGAGACTTGAAGTTGCACTCAAGGAGGCCCAAGAAAGA 1090
Db 703 AGGGAAGGAAACAGAGAGGTTGAGAGACTTGAATTCGCCCTCAAGGAGGCCCAAGAAAGA 762
QY 1091 GTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGG 1150

Db 763 ATTTAGATTTTGAAGAAGAACCCCAAGGATCGTTCTGAGACTGAGACCCAGACAGAG 822
Qy 1151 AGCAGAGAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGGAA 1210
Db 823 CACAAAGAACAGAGAGAGAGAGAAAGCCAGAGAACTGTTGGAAGTGAAGTGGAA 882
Qy 1211 GCAGTGAAGCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAACTC 1270
Db 883 ATGTTAAACCTTTCAGGTGACACCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAACTC 942
Qy 1271 AGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTCAAGCCCTTGAAGGAAA 1330
Db 943 AGTGAAGCTGAGCTCATGAAGAGAGAGACTTCAAGAAAGTGTCAAGCCCTTGAAGGAAA 1002
Qy 1331 AATTCGCAATTCATCAGAGTTGAATGAAGAAGCAAGAGCTTGTATTACTAAACAAAAG 1390
Db 1003 AATTCGCAACCCCATCAGACTGAATGAAGAAGCAAGAGCTTCTTTATTAATACAAAAG 1062
Qy 1391 TTAGAGCTACAGTGGAAAGCATGTCTATCAGAAATCAAAATGGAAAGGCTTAAACACAG 1450
Db 1063 TTGAGCTCCCAAGTGGAAAGCATGAGATCAGAAATTAATAATGGAGCAAGCCAAAACAG 1122
Qy 1451 GATGAAGAGTCCAAATTAATTAATCTGCTACAGATGACACACAAGCTTCTTCAAGACAT 1510
Db 1123 GAGGAAAGTCCAAATTAATTAATCTGCTACAGTGAACCAACAGCTTCTTCAAGATAC 1182
Qy 1511 AATAATGCAATTTGAAACAAATTTAGGAACCTAAACAGAAAGAGTCAAGAAAGTGCACAGG 1570
Db 1183 AATAATGCACTGAAACAAATTTAGGAACCTGAAAGAGAGAGTCTGAAAGAGTGAATAG 1242
Qy 1571 GCAGTGTGAAGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCTGGCTTCCAAA 1630
Db 1243 GTGGTGTCTGAGGAACCTGAATGGAAAGCTGGAATGGCAGAGAGGCTTGGCTTCCAG 1302
Qy 1631 CAGTGCNAATGGATGAATGAAGCAACCAATTCGCAAGCAGGAGAGGACCTGGAAC 1690
Db 1303 CAGTCCAAATGGATGAATGAAGCAGACCAATTCGCAAGCAGAGAGGACCTGGAAC 1362
Qy 1691 ATGACCATCTCAGGGCTCAGATGAGTCTTACTGTTCTGATTTTCATGCTGAAGAGCA 1750
Db 1363 ATGGCTGTTCTCAGGGCTCAGATGAGTATGCTGATTTCTGATTTTCATGCTGAAGAGCA 1422
Qy 1751 GCGAGAGAGAAATTCATGAGGAAAGAGCAACTGGCAATTCGAGCTGGCAGTTCCTGCTG 1810
Db 1423 GCAAGAGAGAGATTCATGAAGAAAGAGCAACTGGCAATTCGAGCTGGCAGTTCCTGCTG 1482
Qy 1811 AAAGAGATGATGCTTTTGAAGAGAGGAG---GCGAGCAGTCTTGAATGGAGATCAGAGT 1867
Db 1483 AAAGCAGCAATGCTTTTGAAGAGAGGAGCAGCAGGCAATCTTGAATGGAGATCAGAGC 1542
Qy 1868 CGTCATGGGCGAGAACAGTGAATCTGACAGCAGGCTTACCTTGTTCAGAGAGGAGCT 1927
Db 1543 CGTCATGGGCGAGAACAGTGAATCTGACAGCAGGCTTTCCTTGTTCAGAGAGGAGCT 1602
Qy 1928 GAGGACAGGAGTGGCGCAACAG-----CGGAATATTCGATTCATTCCTGCCCCAAG 1981
Db 1603 GAGGATAGAACTGGCTGCAACACACACAGAAATTTCAATTCATTCCTGCCCCAAG 1662
Qy 1982 TGTGGAGAGTTCGCTGACATAGACAGTTACAGATTCACGTGATGATTCATTCATTCATTC 2041
Db 1663 TGTGGAGAGTTCGCTGACATAGATACATCTGATTCAGCTTACGAGTTCATTCATTCATTC 1722
Qy 2042 TAAGTGTGATGATATCACTCTCCCCAAA 2068
Db 1723 TAAGTGTGATTTTACTTCCCCAAA 1749

RESULT 27
AD043191
ID AD043191 standard; cDNA; 2023 BP.
XX
AC AD043191;

XX DT 29-JUL-2004 (first entry)
XX DE Mouse optineurin cDNA.
XX KW Mouse; optineurin; glaucoma; diagnosis; gene therapy; ophthalmological;
XX KW gene; ss.
XX OS Mus musculus.
XX FH Key Location/Qualifiers
XX FT CDS 151..1905
XX FT /*tag= a
XX FT /product= "Mouse optineurin"
XX PN WO2004039312-A2.
XX PD 13-MAY-2004.
XX PF 26-JUN-2003; 2003WO-US020165.
XX PR 25-OCT-2002; 2002US-00281457.
XX PA (UYCO-) UNIV CONNECTICUT.
XX PA (SGEO-) ST GEORGES ENTERPRISES LTD.
XX PI Sarfarazi M, Rezaie T, Child AH;
XX DR WPI; 2004-376046/35.
XX DR P-ESDB; AD043192.
XX PT New optineurin nucleic acids and polypeptides having sequence
XX PT alterations, which indicates the presence of an optineurin-associated
XX PT (risk of) glaucoma, useful for detecting or diagnosing glaucoma, e.g.
XX PT primary open angle glaucoma.
XX PS Example; SEQ ID NO 9; 110pp; English.
XX CC The present sequence is that of mouse optineurin (optic neuropathy
XX CC inducing protein) cDNA. It encodes a 584-amino acid proteins that shows
XX CC 78% identity to human optineurin AD043184. Comparison of protein
XX CC sequences, and Western blotting results, indicate conservation of
XX CC optineurin as shown in human, macaque, mouse, rat, pig and cattle
XX CC sequences. Overall, human optineurin has 78-85% sequence identity with
XX CC its homologues in mouse, rat, pig and cattle, and 96% identity with
XX CC macaque. Mutation of the human optineurin gene is associated with primary
XX CC open angle glaucoma. 2 Mutations, E50K and M98K, found in human
XX CC optineurin are also found in macaque, corroborating evidence that this
XX CC mutation is a risk factor for glaucoma. The E50K mutation is further
XX CC conserved in mouse and bovine. Methods of detection, prognosis and
XX CC diagnosis of the presence or absence of optineurin-associated glaucoma or
XX CC of an optineurin-associated increased risk of glaucoma are described, in
XX CC which a sample is tested for the presence of certain gene sequence
XX CC alterations or for alterations in the expression or activity of the
XX CC optineurin protein. Also described are methods of therapy of glaucoma,
XX CC including gene therapy methods.
SQ Sequence 2023 BP; 614 A; 473 C; 578 G; 358 T; 0 U; 0 Other;

Query Match 53.0%; Score 1101.8; DB 12; Length 2023;
Best Local Similarity 77.5%; Pred. No. 2.4e-288;
Matches 1439; Conservative 0; Mismatches 352; Indels 66; Gaps 6;
Qy 263 GAGAAAGTGGCAACTTTTGGAGTGAATTTTCCACAGGAACTTTGCAATGTCCCAATCAA 322
Db 103 GAAGCAGTGACCTAAGAGGAGGAGATCCCTGTGGCGGACCTGTACCATGTCCCATCAA 162
Qy 323 CCTCTCAGCTGCTCACTGAAGAGGAGGAGAGCCCGCCAGTGAAGCAGAGAAATGACCC 382
Db 163 CCTCTCAGCTGCTCACTGAAGAGGAGGAGAGCCCGCCAGTGAAGCAGAGAAATGACCC 222
Qy 383 CCCACCTGGCCACCCAAACCTGGACACGTTTACCCCGGAGGAGCTGCTGCGAGCAGATG 442

QY 1369 GCTTGTATTACTAAACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAA 1428
DB |||||||
DB 382 GCTTGTATTACTAAACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAA 441
QY 1429 AATGGAACAGGCTTAAACACAGAGATGAAAAGTCCAAATTAACCTGTGCTACAGATGACACA 1488
DB |||||||
DB 442 AATGGAACAGGCTTAAACACAGAGATGAAAAGTCCAAATTAACCTGTGCTACAGATGACACA 501
QY 1489 CAACAAGCTTCTTCAAGAACATAATATGATTCATTGAAAAACAATTGAGGAACCTACAAAGAAA 1548
DB |||||||
DB 502 CAACAAGCTTCTTCAAGAACATAATATGATTCATTGAAAAACAATTGAGGAACCTACAAAGAAA 561
QY 1549 AGAGTCAGAAAAAGTGGACAGGCGAGTGTCTGAAGGAACTGAGTGAAGAACTGGAACTGGC 1608
DB |||||||
DB 562 AGAGTCAGAAAAAGTGGACAGGCGAGTGTCTGAAGGAACTGAGTGAAGAACTGGAACTGGC 621
QY 1609 AGAAGAGGCTCTGGCTTCCAAAACAGCTGCAATGGATGAAATGAAGCAACCACTTCCCAA 1668
DB |||||||
DB 622 AGAAGAGGCTCTGGCTTCCAAAACAGCTGCAATGGATGAAATGAAGCAACCACTTCCCAA 681
QY 1669 GCAGGAAGGACCTGGAAACCATGACCAT 1698
DB |||||||
DB 682 GCAGGAAGAGG-CTTGGAAACCATGACCAT 710

RESULT 31

ABQ56229
ID ABQ56229 standard; cDNA; 779 BP.

XX AC ABQ56229;

XX XX
DT 22-AUG-2002 (first entry)XX XX
DE Human ovarian antigen HVUG31 cDNA, SEQ ID NO:2109.

KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; gene; ss.

XX Homo sapiens.

XX OS
XX WO200200677-A1.

XX PN

XX XX
XX 03-JAN-2002.XX XX
XX 07-JUN-2001; 2001WO-US018569.XX PF
XX XX
XX 07-JUN-2000; 2000US-0209467P.XX PR
XX XX
XX (HUMA-) HUMAN GENOME SCI INC.XX PA
XX Birse CE, Rosen CA;XX PI
XX WPI; 2002-147878/19.XX DR
XX P-PSDB; ABP43152.XX XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,

PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian

PT cancer), immune disorders, cardiovascular disorders and neurological

PT diseases.

XX PT
XX Claim 1; SEQ ID NO 2109; 2922pp; English.PS The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also

CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prophesying or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), and toxic
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX XX

SQ Sequence 779 BP; 256 A; 167 C; 221 G; 129 T; 0 U; 6 Other;

Query Match 31.0%; Score 644.4; DB 6; Length 779;
Best Local Similarity 93.5%; Pred. No. 3.2e-164;
Matches 732; Conservative 0; Mismatches 26; Indels 25; Gaps 5;

QY 342 ABAAGGAGGACAGCCCCAGTGAAGCACAGAGAAATGGACCCCCACCTGGCCACCCCAA 401

DB |||||||
DB 6 AAGNTTGGACTGCCCCAGTGAAGCACAGAGAAATGGACCCCCACCTGGCCACCCCAA 65

QY 402 ACCTGGACACGTTTATCCCGGAGGAGCTGCTGCGAGCAGATGAAAGAGCTCTTGACCAAGA 461

DB |||||||
DB 66 ACCTGGACACATTTATCCCGGAGGAGCTGCTGCGAGCAGATGAAAGAGCTCTTGACCAAGA 125

QY 462 ACCACAGCTGAAGAGCCATGAAGTAAATCAAGCCATGAAGGAGGAGATTGAGG 521

DB |||||||
DB 126 ACCACAGCTGAAGAGCCATGAAGTAAATCAAGCCATGAAGGAGGAGATTGAGG 185

QY 522 AGCTTTGGCCTGGACAGAGAAACAGAGAAAGCCAGCTTTTGTAGATACAGAGCA 581

DB |||||||
DB 186 AGCTTTGGCCTGGACAGAGAAACAGAGAAAGCCAGCTTTTGTAGATACAGAGCA 245

QY 582 AAGAAGCAAAAGAGCGTCTAATGGCCTTGAGTCATGAGAAATGAGAAATTGAAGGAAGAGC 641

DB |||||||
DB 246 AAGAAGCAAAAGAGCGTCTAATGGCCTTGAGTCATGAGAAATGAGAAATTGAAGGAAGAGC 305

QY 642 TTGGAATACTAAAGGGAATCAGAAAGTCTCTGAGGACCCCACTGATGCTCCAGGC 701

DB |||||||
DB 306 TTGGAATACT-AAAAGGGAATCAGAAAGTCTCTGAGGACCCCACTGATGCTCCAGGC 364

QY 702 TTCCAGGGCCGAGGAGAGGAAAGGACAGCTCAGGACCCAGGCTGGTGGAGGCTAC 761

DB |||||||
DB 365 TTCCAGGGCCGAGGAGAGGAAAGGACAGCTCAGGACCCAGGCTGGTGGAGGCTAC 424

QY 762 AAGCAGAGAAGGAGACCTGTTGGGCATCGTGTCTGAACCTGACGCTCAAGCTGAACTCCA 821

DB |||||||
DB 425 AAGCAGAGAAGGAGACCTGTTGGGCATCGTGTCTGAACCTGACGCTCAAGCTGAACTCCA 484

QY 822 GCGGCTCTCAGAAAGATTCTTTGTTGAAATTTAGGATGGCTGAAGGAGAGCAGAGGCT 881

DB |||||||
DB 485 GCGGCTCTCAGAAAGATTCTTTGTTGAAATTTAGGATGGCTGAAGGAGAGCAGAGGCT 544

QY 882 CAGTAAAGAAATCAAGCATAGTCTCTGGGCCACGAGAACAGCTCTCCACTGGCAGCGCAT 941

DB |||||||
DB 545 CAGTAAAGAAATCAAGCATAGTCTCTGGGCCACGAGAACAGCTCTCCACTGGCAGCGCAT 998

Db 1023 CCGAGGACAGGAGCTGGCAGCAGCGGGCGCAGAGCCTCGCAGTATTCGATTCACTCTCTGCC 1082

Qy 1977 CCAAGTGTGGAGAGTTCCTGCTGACATAGACACGTTACAGATTTCACTGATGGATGCA 2036

Db 1083 CCAAGTGGGGAGGTCCTGCGGACATCGACAGCTTCAGATCCATGTGATGGACTGCA 1142

Qy 2037 TCATTTAAGTGTGATGATACCTCCCAAACTGTTGGT 2077

Db 1143 TCATTTAGTGT---TCTCTCAGTCCCCAAAGCTCTTGGT 1180

RESULT 33

ADL24739

ID ADL24739 standard; DNA; 815 BP.

AC ADL24739;

XX

XX 20-MAY-2004 (first entry)

XX Intestinal epithelium/peyer's patch M cell-associated DNA sequence #75.

XX

XX intestinal epithelium cell development; peyer's patch M cell development;

XX inflammatory bowel disease; glutenenteropathy; infectious disease;

XX autoimmune disease; haemolytic anaemia; rheumatoid arthritis; dermatitis;

XX Grave's disease; multiple sclerosis; allergy; asthma; diabetic mellitus;

XX immune system disorder; hypersensitivity; anaphylaxis;

XX blood group incompatibility; ds; mouse; murine.

XX

OS Mus musculus.

XX

XX WO200280852-A2.

XX

XX 17-OCT-2002.

XX

XX 04-APR-2002; 2002WO-US010873.

XX

XX 04-APR-2001; 2001US-0281416P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX

XX Brayden DJ, Byrne D, O'mahony DJ, Evans CF, Mah SP, Lo DD;

PI WPI; 2003-075470/07.

XX

XX Novel isolated or purified polypeptide encoded by genes associated with

PT intestinal epithelium or M cell development, differentiation or function,

PT useful for treating autoimmune diseases and infectious diseases.

XX

XX Claim 1; SEQ ID NO 249; 152pp; English.

XX

XX The invention comprises DNA sequences which are associated with

CC intestinal epithelium and peyer's patch M cells. The DNA sequences of the

CC invention are useful for assessing, modifying, modulating or regulating

CC intestinal epithelium or M cell development. The DNA sequences of the

CC invention are also useful in the treatment of: inflammatory bowel

CC disease, glutenenteropathy, infectious diseases, autoimmune diseases

CC (e.g. haemolytic anaemia, rheumatoid arthritis, dermatitis, Grave's

CC disease, multiple sclerosis, allergy, asthma and diabetic mellitus),

CC diseases or disorders of the immune system, hypersensitivity,

CC anaphylaxis, and blood group incompatibility. The present nucleic acid

CC represents an intestinal epithelium/peyer's patch M cell-associated DNA

CC sequence of the invention. NOTE: The present sequence is not shown in the

CC specification, but has been retrieved from the WIPO website.

XX

XX Sequence 815 BP; 259 A; 195 C; 238 G; 133 T; 0 U; 0 Other;

XX

XX

Query Match 26.2%; Score 544; DB 10; Length 815;

Best Local Similarity 81.9%; Pred. No. 6.5e-137;

Matches 668; Conservative 0; Mismatches 135; Indels 13; Gaps 3;

Qy 1213 ACTGAACCTCCAGTGCATCTCTGTTTAAAGAGCTTCAAGAGCTCATACAAACTCAG 1272

Db 1 ACTGAGCATCAAGTGACCTCTCTGTTTAAAGAGCTTCAAGAGGCACACAAACTCAG 60

Qy 1273 CGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAGTGTGAGGCCCTTTGAAAGGAAAAA 1332

Db 61 TGAGCTGAGCTGATGAAGAAGAGACTTCAAGAAAAGTGTGAGGCCCTTTGAGAGGAAGA 120

Qy 1333 TTCTGCAATTCATCAGAGTGAATGAAGAAGCAAGAGCTTGTATTACTTAAACAAAAGTT 1392

Db 121 CTCTGCAACACCATCAGAGCTGAATGAAGAAGCAAGAGCTGTTTACAGTAACAGAAGTT 180

Qy 1393 AGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAACAGGCTTAAACAGAGA 1452

Db 181 AGAGCTGAGGTGGAGAGCATGCGCTCCGAATCAAGATGAGGAGGCCCAAGACAGAGA 240

Qy 1453 TGAAAAATCCAAATTAATCTGTGCTACAGATGACACACAAAGAGCTTCTTCAAGAAATAA 1512

Db 241 GGAGAAATCCAGGTTAGCCACTCTGCGAGCAACTCACAACAAGCTCTTCAAGAAATAA 300

Qy 1513 TAATGCATTGAAACCAATTTGAGGAATCAACAGAAAGAGTTCAGAAAAAGTGGACAGGGC 1572

Db 301 TAAGGCATGAAAAACAATTTGAAGAACTAACAGCAACAGGCAGAAAAGTGGACAAGAT 360

Qy 1573 AGTGTCTGAAGGAACCTGAGTGAATAAACTGGAACTGGCAGAGAAGGCTCTGGCTTCCAAACA 1632

Db 361 GTTGTCTGAGGAGCTCAGCGAGAAGCTGGAGCTGGCAGAGCAGGCTCTGGCATCCAAACA 420

Qy 1633 GCTGCAAAATGGATGAATAAGAACAAACATTTCCAAAGCAGGAGGAGACCTGGAACCAT 1692

Db 421 GCTCCAGATGGATGAGATGAAGCAGACGCTCGCTAAGCAGGAGGAAGACCTGAGAGACCA 480

Qy 1693 GACCATCTCAGGGCTCAGATGGAAAGTTTACCTGTTCTGATTTTTCATGCTGAAAGACAGC 1752

Db 481 GCGCGCTCTCAGGGCTCAGATGGAGGTGACTGCTCAGATTTTCAGCTGAGAGAGCAGC 540

Qy 1753 GAGAGAGAAAATTCATGAGGAAAAGGAGCAACTGGCAATTCGAGCTGGCAGTTCTGCTGAA 1812

Db 541 AAGAGAGAAGATTTCATGAG- AAGAGAGCAGCTGGCTTTCAGCTGCGGATTTGCTGAA 599

Qy 1813 AGAGAAATGATGCTTTTCAAGACGAGGC---AGGAGTCTTGTATGAGATGAGAGATGCG 1869

Db 600 AGAGAAACATGACATTTGAAGAGGAGGAGCAGTAGACAGTCCCTGATGAAATTCAGTGGC 659

Qy 1870 TCATGGGGGAGAGCAAGATGACTCTGACCAGCAGGCTTACCTTGTTCAAAGAGGAGCTGA 1929

Db 660 ACACGGGGGAGAGCAAGACAGTACTCTGACCAGCAGACTTACCTGTTTCAAAGAGAGCCGA 719

Qy 1930 GGACAG-----GGACTGGCGGCAACAGCGGAATATTCGATTCATTCTTCGCCCAA 1980

Db 720 GGACAGGAGCTGGCAGCAGCGGAGCAGCTCGCAGTATTCGATTCCTCTGCCCCAA 779

Qy 1981 GTGTGAGAGGTTCTGCTCGCATAGACACAGTTACA 2016

Db 780 GTGCGGGAGGTCCTGCGGAGCATCGACACGCTTCA 815

RESULT 34

AAH57391/C

ID AAH57391 standard; cDNA; 3454 BP.

XX

XX AAH57391;

XX

XX 10-SEP-2001 (first entry)

XX

XX Human skeletal muscle cell specific cDNA sequence SEQ ID NO:231.

XX

XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;

XX liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;

XX metabolic disease; developmental disease; cytostatic; immunomodulatory;

XX neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

XX

XX Homo sapiens.

XX

XX WO200132927-A2.

XX

PD	10-MAY-2001.
XX	
PF	02-NOV-2000; 2000WO-US030396.
XX	
PR	04-NOV-1999; 99US-0163508P.
XX	(INCY-) INCYTE GENOMICS INC.
PA	
XX	Sornasse T, Seilhamer JJ, Watson GA;
PI	WPI; 2001-291057/30.
XX	
DR	New cell and tissue specific polynucleotides useful for diagnosis,
XX	prognosis or monitoring of treatments for disorders where the gene is
PT	associated with a cancer, immunopathology or neuropathology.
PT	
XX	Claim 1; Page 159-160; 327pp; English.
PS	
XX	AH57161 to AH57576 represent cell and tissue specific polynucleotide
CC	sequences (I). (I) can have cytostatic, immunomodulatory and
CC	neuroprotective activities, and can be used in gene therapy. (I) and
CC	proteins (II) encoded by then are used in high throughput screening
CC	assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC	mimetics, peptides, proteins, agonists, antagonists, antibodies or their
CC	fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
CC	agents. Expression of (I) in a sample indicates the differentiation of
CC	embryonic stem cells into a tissue selected from brain, heart, kidney,
CC	liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
CC	to produce an expression profile that defines a metabolic or
CC	developmental process, treatment, condition, disease or disorder. The
CC	gene profile can be used for diagnosis, prognosis or monitoring of
CC	treatments and for investigating a predisposition to a disorder where the
CC	gene is associated with a cancer, immunopathology or neuropathology
XX	
SQ	Sequence 3454 BP; 1094 A; 725 C; 794 G; 841 T; 0 U; 0 Other;
Query Match 25.4%; Score 526.6; DB 4; Length 3454;	
Best Local Similarity 99.1%; Pred. No. 7.5e-132;	
Matches 540; Conservative 0; Mismatches 4; Indels 1; Gaps 1;	
QY	1307 AAGTGTGACGGCCCTTGAAAGGAAAAATTCGCAATCCATCAGAGTTGAATGAAAGCAA 1366
DB	
DB	2549 AGGTGTGACGGCCCTTGAAAGGAAAAATTCGCAATCCATCAGAGTTGAATGAAAGCAA 2490
QY	1367 GAGCTTGTTTTATACTAACAAAAAGTTAGAGCTACAAGTGGAAAAGCATGCTATCAGAAATC 1426
DB	
DB	2489 GAGCTTGTTTTACTTAAC-AAAGTTAGAGCTACAAGTGGAAAAGCATGCTATCAGAAATC 2431
QY	1427 AAAATGAAACAGGCTAAAAACAGAGGATGAAAAGTCCAAATTTAAGTGTCTACAGATGACA 1486
DB	
DB	2430 AAAATGAAACAGGCTAAAAACAGAGGATGAAAAGTCCAAATTTAAGTGTCTACAGATGACA 2371
QY	1487 CACAACNAGCTTTTCAAGAACAATAATATGCATTGAAAACAATTTAGGGAACCTAACCAAGA 1546
DB	
DB	2370 CACAACNAGCTTTTCAAGAACAATAATATGCATTGAAAACAATTTAGGGAACCTAACCAAGA 2311
QY	1547 AAAGAGTCAAAAAAGTGGACAGGCGAGTCTGTAAGGAACTGAGTGGAAAACCTGGAACCTG 1606
DB	
DB	2310 AAAGAGTCAAAAAAGTGGACAGGCGAGTCTGTAAGGAACTGAGTGGAAAACCTGGAACCTG 2251
QY	1607 GCAGAGAAGGCTCTGGCTTCCAAACAGCTGCCAAATGGATGAAATGAAAGCAAAACCATTGCC 1666
DB	
DB	2250 GCAGAGAAGGCTCTGGCTTCCAAACAGCTGCCAAATGGATGAAATGAAAGCAAAACCATTGCC 2191
QY	1667 AACCAGGAGAGACCTGGAAAAACATGACCATCTCTCAGGGCTCAGATGGAGTTTACTGT 1726
DB	
DB	2190 AACCAGGAGAGACCTGGAAAAACATGACCATCTCTCAGGGCTCAGATGGAGTTTACTGT 2131
QY	1727 TCTGATTTTTCATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGAGGAAAAGGAGCAACTG 1786
DB	
DB	2130 TCTGATTTTTCATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGAGGAAAAGGAGCAACTG 2071
QY	1787 GCATTGAGCTGGCAGTTCTTGCTGAAAGAGAATGATGCTTTTCGAAGACGGAGGCAGGCAG 1846

CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library
SQ Sequence 685 BP; 259 A; 114 C; 172 G; 132 T; 0 U; 8 Other;

Query Match 24.7%; Score 513.4; DB 10; Length 685;
Best Local Similarity 98.1%; Pred. No. 1.2e-128;
Matches 561; Conservative 0; Mismatches 7; Indels 4; Gaps 4;
QY 956 AGCAGATCTGCAGATGGGCGCCAAAGAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAG 1015
|||
Db 7 AGCGGATCTGCAGATGGGCGCCAAAGAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAG 66
QY 1016 CTCCTGCTGCTTAAGGGAAGGGAATCAGAAGGTGGAGACTTCAAGTTGCACTCAAG 1075
|||
Db 67 CTCCTGCTGCTTAAGGGAAGGGAATCAGAAGGTGGAGACTTCAAGTTGCACTCAAG 126
QY 1076 GAGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAAATCGTTCTGAGATTGAA 1135
|||
Db 127 GAGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAAATCGTTCTGAGATTGAA 186
QY 1136 ACCCAGACAGAGGGGAGCAGAGAAAGAAATGATGAAGAGAACGCCCGGAGACTGTT 1195
|||
Db 187 ACCCAGACAGAGGGGAGCAGAGAAAGAAATGATGAAGAGAACGCCCGGAGACTGTT 246
QY 1196 GGAAGCGAAGTGAAGCACTGACCTCCAGTGCACATCTCTGTTTAAGGAGCTTCAAGAG 1255
|||
Db 247 GGAAGCGAAGTGAAGCACTGACCTCCAGTGCACATCTCTGTTTAAGGAGCTTCAAGAG 306
QY 1256 GCTCATACAAACTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAGTGTCTAG 1315
|||
Db 307 GCTCATACAAACTCAGCGAAGCTGGA-CTAATGAAGAAGAGACTTCAAGAAAGTGTCTAG 365
QY 1316 GCCTTTGAAAGGAAAATTC-TGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGT 1374
|||
Db 366 GCCTTTGAAAGGAAAATTC-TGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGT 425
QY 1375 TTATACTAACAAAAAGTTAGAGCTACAAGTGGAAAGCATCTATCAGAAATCAAAATGGA 1434
|||
Db 426 TTATACTAACAAAAAGTTAGAGCTACAAGTGGAAAGCATCTATCAGAAATCAAAATGGA 485
QY 1435 ACAGGCTTAAACAGAGGATGAAAGTCCAAATTAACCTGTCTACAGATGACACACAACAA 1494
|||
Db 486 ACA-GCTAAGCAGAGGATGAAAGTCCAAATTAACCTG-NCTACAGATGACACACAACAA 543
QY 1495 GCTTCTTCAAGAACATAAATATGCAATTGAAAA 1526
|||
Db 544 GCTTCTTCAAGAACATAAATATGCAATTGAAACA 575

RESULT 36
ADF79591
ID ADF79591 standard; DNA; 458 BP.
XX
AC ADF79591;
XX

DT 26-FEB-2004 (first entry)
XX
DE Leukaemia-related DNA sequence #147.
XX
KW Cytostatic; Gene therapy; leukaemia; ss.
XX
OS Unidentified.
XX
FN W02003039443-A2.

XX 15-MAY-2003.
PD
XX
XX
PF 04-NOV-2002; 2002WO-EP012303.
XX
XX
PR 05-NOV-2001; 2001EP-00126244.
PR 30-APR-2002; 2002EP-00009758.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (UYLU-) UNIV LUDWIG MAXIMILIANS.
PA (HAFE/) HAFERLACH T.
PA (SCHO/) SCHOCH C.
PA (KERN/) KERN W.
XX
XX Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;
PI Bils R, Brors B, Mergenthaler S;
XX
XX WPI; 2003-505037/47.
DR
XX
XX
PT Determining the subtype of leukemia cells and whether a patient sample
PT contains leukemia cells or other cells, useful for treating leukemia,
PT comprises determining the expression profile of a group of markers in a
PT patient sample.
XX
XX Disclosure; SEQ ID NO 147; 2938pp; English.
PS
XX
XX The present invention relates to a method (M1) for determining the
CC subtype of leukaemia cells and whether a patient sample contains
CC leukaemia cells. The method comprises determining the expression profile
CC of a group of markers in a patient sample. The method is useful for
CC determining the presence of leukaemia cells, its types or subtypes, and
CC for the preparation of a medicament for treating leukaemia.
XX
SQ Sequence 458 BP; 135 A; 95 C; 130 G; 98 T; 0 U; 0 Other;

Query Match 22.1%; Score 458; DB 10; Length 458;
Best Local Similarity 100.0%; Pred. No. 1.2e-113;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1609 AGAGAAGGCTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCAATGCCAA 1668
|||
Db 1 AGAGAAGGCTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCAATGCCAA 60
QY 1669 GCAGGAAGAGGACCTGGAAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTC 1728
|||
Db 61 GCAGGAAGAGGACCTGGAAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTC 120
QY 1729 TGATTTTTCATGCTGAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGAGCAACTGGC 1788
|||
Db 121 TGATTTTTCATGCTGAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGAGCAACTGGC 180
QY 1789 ATTGAGCTGGCAGTTCTCTGTAAGAGAGAAATGATGCTTTGAGAGCGGAGGAGGAGTC 1848
|||
Db 181 ATTGAGCTGGCAGTTCTCTGTAAGAGAGAAATGATGCTTTGAGAGCGGAGGAGGAGTC 240
QY 1849 CTTGATGAGAGTCGAGAGTCGTCATGGGCGAGAAACAAGTACTCTGACAGAGGCTTGA 1908
|||
Db 241 CTTGATGAGAGTCGAGAGTCGTCATGGGCGAGAAACAAGTACTCTGACAGAGGCTTGA 300
QY 1909 CCTTGTTCAAAGAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTCGGATTCA 1968
|||
Db 301 CCTTGTTCAAAGAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTCGGATTCA 360
QY 1969 TTCCTGCCCAAGTGTGGAGAGGTTCTGCTGACATAGACACGCTTACAGATTACCGTGTAT 2028
|||
Db 361 TTCCTGCCCAAGTGTGGAGAGGTTCTGCTGACATAGACACGCTTACAGATTACCGTGTAT 420
QY 2029 GGATTGCATCAATTAAAGTGTGTGATGTATCACTCCCA 2066
|||
Db 421 GGATTGCATCAATTAAAGTGTGTGATGTATCACTCCCA 458

QY 956 AGCAGATCTGAGATGGGGCCAGAAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAG 1015
 DB 1 AGCAAAATCTGCAGAGGTACCAGGAATTATGTGGAGTTTGAGGAATTAATCTGTGAGCCAG 60
 QY 1016 CTCCTGCTGTGCTTAAGGAGGGAATCAGAACTGAGAGACTTGAAGTTGCACTCAAG 1075
 DB 61 CTCCTGCTGTGCTTAAGGAGGGAATCAGAACTGAGAGACTTGAAGTTGCACTCAAG 120
 QY 1076 GAGGCCAAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAAAGTAACTGTTCTGAGATTGAA 1135
 DB 121 GAAGCCAAAGAAAGAAATTTCTGATTTTGAAGAAAGAAACCAAGGATCAATCTGAGACTGAG 180
 QY 1136 ACCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTT 1195
 DB 181 ACCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTT 240
 QY 1196 GGAAGCCAAAGTGAAGCACTGACCTCCAGTGCATCTGTTTAAAGGAGCTTCAAGAG 1255
 DB 241 GGAAGTGAAGTGAAGCACTGACCTCCAGTGCATCTGTTTAAAGGAGCTTCAAGAG 300
 QY 1256 GCTCATACAAAACCTCAGCGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTTCAG 1315
 DB 301 GCTCACAGAACTCAGTGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTTCAG 360
 QY 1316 GCCCTTGAAGGAAAAATTTGCAATTTCCATCAGAGTTGAATGAAAAAGCAAGAGCTTGT 1375
 DB 361 GCCCTTGAAGGAAAAATTTGCAAGGCCATCAGAACTGAATGAAAAAGCAAGAGCTGTT 420
 QY 1376 TATACTTAACAAAAG 1390
 DB 421 TATATAAACAATAAG 435

RESULT 40

ABX49269
 ID ABX49269 standard; cdNA; 416 BP.
 XX
 AC ABX49269;
 XX
 DT 21-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #14434.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 XX 24-SEP-2001; 2001US-00960352.
 XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 DR WPI; 2003-110599/10.
 XX
 PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 PS Claim 2; SEQ ID NO 14434; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived from
 CC cattle, and the LMPD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 1512 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 1512 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridization between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMPD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 1512 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 SQ Sequence 416 BP; 167 A; 75 C; 102 G; 72 T; 0 U; 0 Other;

Query Match 16.1%; Score 333.4; DB 8; Length 416;
 Best Local Similarity 87.7%; Pred. No. 8.1e-80;
 Matches 364; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 1001 TTAACCTGTGAGCCAGCTCTCTGCTGCTTAAGGAAAGGGAATCAGAGGTGGAGAGACTT 1060
 DB 2 TTAACCTGTGAGCCAGCTCTCTGCTGCTTAAGGAAAGGGAATCAGAGGTGGAGAGACTT 61
 QY 1061 GAAGTTTGCACTCAAGGAGGCCAAAGAAAGAGTTTTCAGATTTTCAAAAAGAAACAAGTAAT 1120
 DB 62 GAAGTTTGCCCTCAAGGAAGCCAAAGAAAGAAATTTCTGATTTTGAAGAAGCAAGGAT 121
 QY 1121 CGTTCCTGAGATTGAAACCCAGACAGAGGGGAGCACAGAGAAAGAGAAATGATGAAGAGAAA 1180
 DB 122 CATTCCTGAGACTGAGACCCAGACAGAGAGGAGCACACAGAACAGAGAAAGAGAGAGAAA 181
 QY 1181 GCGCCGGAGACTGTTGGAAAGCGAAGTGGAGAGCTGAACTCCAGGTGACATCTCTGTTT 1240
 DB 182 GACACCAAAAACCTATTGGAAAGTGAAGTGGAAACACACTGAACCTTCAGGTGACCAACCTGTTT 241
 QY 1241 AAGGAGCTTCAAGAGGCTCATACAAAACCTCAGCGAAGCTGAGCTAATGAAGAGAGACTT 1300
 DB 242 AAGGAGCTTCAAGAGGCTCACAGAAACCTCAGTGAAGCTGAGCTAATGAAGAGAGACTT 301
 QY 1301 CAAGAAAAGTGTGAGGCTTGAAGGAAAAAATTTCTGCAATTCCTCAGAGTTGAATGAA 1360
 DB 302 CAAGAAAAGTGTGAGGCTTGAAGGAAAAAATTTCTGCAAGCCCATCAGAACTGAATGAA 361
 QY 1361 AAGCAAGAGCTTGTATTATCTAACAAAAAGTTAGAGCTACAAGTGAAGAGCATGTC 1415
 DB 362 AAACAGAGAGCTGGTTTATATAAAAAACAAAAAGTTAGAGCTCCAAAGTGGAAAGCATGC 416

RESULT 41

AED67129
 ID AED67129 standard; DNA; 364 BP.
 XX
 AC AED67129;
 XX
 DT 12-JAN-2006 (first entry)
 XX
 XX Human marker DNA sequence SEQ ID NO:55.
 XX

KW ds; selectable marker; biosensor; gynecological disorder; gynecological;
KW genitourinary disease; gynecology and obstetrics; immunoassay;
KW endometriosis; ovary tumor; cytostatic; endocrine disease; neoplasm;
KW breast tumor; uterine cervix tumor; cancer.
XX Homo sapiens.
XX US2005239146-A1.
XX 27-OCT-2005.
XX 17-FEB-2005; 2005US-00060867.
XX 04-DEC-2001; 2001US-00004587.
XX (TAIN/) TAINSKY M.
XX (DRAG/) DRAGHICI S.
XX (CHAT/) CHATTERJEE M.
XX Tainsky M, Draghici S, Chatterjee M;
XX WPI; 2005-757147/77.
XX P-PSDB; AED67130.
XX New biosensor comprises detection means for detecting a presence of at
PT least one marker indicative of a specific disease, useful for detecting
PT the presence of diseases or diagnosing diseases, e.g. gynecological
PT disease.
XX Disclosure; SEQ ID NO 55; 119pp; English.
XX The invention relates to a novel biosensor for use in detecting the
CC presence of diseases, comprising a detector for detecting a presence of
CC at least one marker indicative of a specific disease. The disease
CC detected by the biosensor is a gynecological illness. The detection means
CC is selected from an assay, a microarray, a macroarray, a slide, or a
CC filter containing specific biomarkers of disease. Preferably, the disease is a
CC gynecological disease selected from endometriosis, ovarian cancer, breast
CC cancer, cervical cancer, or primary peritoneal carcinoma. The targets are
CC personalized to the individual receiving treatment. The biosensor is
CC useful as for detecting the efficacy of a pharmaceutical. It is also
CC useful as staging means for detecting the disease stage. The biosensor is
CC useful for detecting the presence of diseases including gynecological
CC diseases, e.g. endometriosis, ovarian cancer, breast cancer, cervical
CC cancer, or primary peritoneal carcinoma. It can also be used for
CC diagnosing cancer. The present sequence represents a marker DNA sequence
CC that may be used in a biosensor of the invention.
XX SQ Sequence 364 BP; 131 A; 52 C; 89 G; 67 T; 0 U; 25 Other;
Query Match 15.9%; Score 330; DB 14; Length 364;
Best Local Similarity 93.0%; Pred. No. 6.4e-79;
Matches 330; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 979 GAATTACTTCGAACATGAGGAGTTAACTGTGACCCAGCTCTCTGTGCTTAAGGAAGG 1038
DB 10 GAATTACTTCGAACATGAGGAGTTAACTGTGACCCAGCTCTCTGTGCTTAAGGAAGG 69
QY 1039 GAATCAGAGCTGGAGAGCTTCAAGTTGCATCAGAGGCCCCAAGAAAGAGTTTCAGA 1098
DB 70 GAATCAGAGCTGGAGAGCTTCAAGTTGCATCAGAGGCCCCAAGAAAGAGTTTCAGA 129
QY 1099 TTTTGAAGAAAGAAACAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGACACAGA 1158
DB 130 TTTTGAAGAAAGAAACAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGACACAGA 189
QY 1159 GAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGAAGCACTGAA 1218
DB 190 GAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGAAGCACTGAA 249
QY 1219 CCTCCAGGTGACATCTCTGTTTAAGAGCTTCAAGAGGCTCATACAAAACCTCAGCGAAGC 1278

Db 250 CCTNCAGGTGACATNTNTGTTTAAGGAGCTTNAANAGGCTNATACAAAACCTNANCAANC 309
QY 1279 TGAGCTAATGAAGAAGAGACTTCAAGAAAAGTGTCAAGCCCTTGAAGGAAAAAT 1333
DB 310 TGANCTAATGAANAANANANACTTNAANAANAANTGTNANGCCNTTGAANGNNAAT 364
RESULT 42
ADU14099
ID ADU14099 standard; DNA; 493 BP.
XX AC ADU14099;
XX 27-JAN-2005 (first entry)
XX Solid tumour prognosis gene seqid 4538.
XX cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX Unidentified.
XX WO2004097052-A2.
XX 11-NOV-2004.
XX 29-APR-2004; 2004WO-US013587.
XX 29-APR-2003; 2003US-0466067P.
XX 23-JAN-2004; 2004US-0538246P.
XX (AMHP) WYETH.
XX (STRA/) STRAHS A.
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
PI Immerman F, Dörner AJ;
PI WPI; 2004-804779/79.
XX A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX Disclosure; Page; 11pp; English.
XX The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially
CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumour, and
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.

CC The method, system, and array are useful for prognosing and treating
CC solid tumours. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIFO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 493 BP; 101 A; 74 C; 97 G; 72 T; 0 U; 149 Other;

Query Match 15.0%; Score 312; DB 13; Length 493;
Best Local Similarity 69.2%; Pred. No. 5.9e-74;
Matches 334; Conservative 0; Mismatches 147; Indels 2; Gaps 2;

QY 1597 ACTGGAAGTGCAGAGAGGCTCTGGCTTCCAAACA-GCTCAATGATGAATGAAGC 1655
Db 1 ACTGGAAGTGCAGAGAGGCTCTGGCTTCCAAACANNNNNNNNNNNNNNNNNNN 60

QY 1656 AAACCATTCGCCAGCAGAGAGGACCTGGAAACCATGACCAT-CCTCAGGGCTCAGATG 1714
Db 61 NNN 120

QY 1715 GAAGTTTACTGTTCTGATTTTTCATGCTGAAGAGCAGCAGAGAGAAATTCATGAGAA 1774
Db 121 NNN 180

QY 1775 AAGCAGCAACTGGCATTCGACCTGGCAGTTCCTGCTGAAGAGATGATGCTTTCGAAGAC 1834
Db 181 AAGCAGCAACTGGCATTCGACCTGGCAGTTCCTGCTGAAGAGATGATGCTTTCGAAGAC 240

QY 1835 GGAGCAGGAGCTCTTGATGGAGATGCAGAGTCGTATGGGGCGAGACAAGTGAATCT 1894
Db 241 NNAGCAGGAGCTCTTGATGGAGATGCAGAGTCGTATGGGGCGAGACAAGTGAATCT 300

QY 1895 GACCAGCAGGCTTACCTTGTTCAGAGAGGAGCTCAGGACAGGGAGCTGGCGCAACAGCGG 1954
Db 301 GACCAGCAGGCTTACCTTGTTCAGAGAGGAGCTCAGGACAGGGAGCTGGCGCAACAGCGG 360

QY 1955 AATATTCGATTCATTCCTGCCCCAAGTGTGGAGAGGTTCTGCTGATAGACACGTTA 2014
Db 361 AATATTCGATTCATTCCTGCCCCAAGTGTGGAGAGGTTCTGCTGATAGACACGTTA 420

QY 2015 CAGATTCAGTGTGATTCATTCATTTAAGTGTGATGATACCTCCCAAACTGTT 2074
Db 421 CAGATTCAGTGTGATTCATTCATTTAAGTGTGATGATACCTCCCAAACTGTT 480

QY 2075 GGT 2077
Db 481 GGT 483

RESULT 43
AEB53489
ID AEB53489 standard; DNA; 493 BP.
XX
AC AEB53489;
XX
DT 06-OCT-2005 (first entry)
XX
DE Human probe for epithelial cell DNA chip SEQ ID 448.
XX
KW Probe; ss; endothelial cell; gene expression; DNA chip; Antiinflammatory;
KW Vasotropic; Gastrointestinal-Gen.; Cytostatic; Immunomodulator;
KW Immunosuppressive; Antiarteriosclerotic; Neuroprotective; Nootropic;
KW Ophthalmological; CNS-Gen; Gynecological; Hepatotropic; Antipsoriatic;
KW Respiratory-Gen; Antiasthmatic; Antiallergic; Dermatological;
KW Antibacterial; Respiratory-Gen.; Inflammation; atherosclerosis;
KW transplant rejection; tumor; inflammatory bowel disease; psoriasis;
KW reperfusion injury; respiratory distress syndrome; asthma;
KW allergic rhinitis; dermatitis; bacterial meningitis; encephalitis;
KW uveitis; leucocyte disorder; central nervous system disorders;
KW Alzheimers disease; endometriosis; multiple sclerosis;
KW alcoholic hepatitis; bacterial pneumonia; lung inflammation; pleurisy;
KW chronic bronchitis; bronchiectasis; cystic fibrosis;
KW chronic obstructive pulmonary disease; vasculitis; polyarteritis nodosa;

extracranial temporal arteritis; pre-eclampsia; autoimmune diseases.
Homo sapiens.
WO2005068655-A2.
28-JUL-2005.
14-JAN-2005; 2005WO-GB0000057.
16-JAN-2004; 2004GB-00000976.
(UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.
Smith SK, Charnock-Jones DS, Print CG, Johnson NA;
WPI; 2005-563988/57.
Assessing tissue inflammatory response, by determining level of five transcripts e.g., glycerol kinase, interleukin-6 or proteins encoding transcripts in sample, and comparing abundance of transcripts/proteins with control sample.

Claim 21; SEQ ID NO 448; 492pp; English.

The invention relates to assessing (M1) a tissue inflammatory response, involving determining quantitatively the level of at least five transcripts of Table-1a e.g., glycerol kinase, Tumor necrosis factor (TNF)-inducible A20, insulin-like growth factor binding protein 6, interleukin (IL)-6 or proteins encoded by the transcripts, in a sample, and comparing the abundance of transcripts or proteins with a control sample. Also included are diagnosing (M2) a condition with a tissue inflammatory response (involving determining the abundance of endothelial cell-derived proteins encoded by at least five transcripts of Table-1a in a sample from a patient suspected of suffering from such a condition), a gene chip array (I) for carrying out (M1)/(M2) (comprising at least five nucleic acids suitable for detection of at least five transcripts of Table-1a, optionally a control specific for transcripts, and one or more controls for the gene chip), a gene chip array (II) for detection of at least five transcripts of Table-1a, a protein based assay (M3) suitable for carrying out (M1)/(M2) (for the assessment of at least five proteins encoded by transcripts of Table-1a, optionally a control specific for protein and/or optionally one or more control for the assay), an assay (M4) for determining a modulator of a tissue inflammatory response or their condition, use of a modulator (III) obtained by (M4) for the treatment of tissue inflammatory response or their conditions and a vector (IV) comprising a sequence encoding a transcript from of Table-1b provided in the specification. (M1) is useful for assessing a tissue inflammatory response. The sample comprises cells obtained from a site within a patient affected by a tissue inflammatory response. The cells are endothelial cell, or the sample is of patient blood, serum or urine, where the endothelial cells are human umbilical vein endothelial cells, human coronary artery endothelial cells, or human uterine microvascular endothelial cells. The control sample is obtained from endothelial cells from a tissue of patient, which is not affected by a tissue inflammatory response or obtained from a tissue demonstrating an inflammatory response in a patient at an earlier point in time. (M2) is useful for diagnosing a condition with tissue inflammatory response. The tissue inflammatory response is associated with an inflammatory disease, vasculitic syndrome, atherosclerosis or an associated disease, chronic transplant rejection, where the condition involves tumor growth. The inflammatory disease such as inflammatory bowel disorders, psoriasis, ischemic reperfusion, adult respiratory distress syndrome, asthma, allergic rhinitis, dermatitis, meningitis, encephalitis, uveitis, diseases involving leucocyte diapedesis, central nervous system inflammatory disorders, Alzheimer's, endometriosis, multiple sclerosis, multiple organ injury syndrome, alcoholic hepatitis, bacterial pneumonia, antigen-antibody complex mediated diseases, inflammation of the lung (including pleurisy, alveolitis, pneumonia, chronic bronchitis, bronchiectasis, cystic fibrosis and chronic obstructive pulmonary disease (COPD)), and vasculitis, polyarteritis nodosa, giant cell arteritis, microscopic polyarteritis, pre-eclampsia and autoimmune diseases. (III) is useful for the treatment of tissue inflammatory response or their condition as

CC above. The present sequence is a probe used in the gene chip of the
CC invention to detect a human gene chosen from those appearing in table 1a
CC (known genes).
XX
SQ Sequence 493 BP; 101 A; 74 C; 97 G; 72 T; 0 U; 149 Other;

Query Match 15.0%; Score 312; DB 14; Length 493;
Best Local Similarity 69.2%; Pred. No. 5.9e-74;
Matches 334; Conservative 0; Mismatches 147; Indels 2; Gaps 2;

QY 1597 ACTGGAACCTGGCAGAGAGGCTCTGGCTTCCAAACA-GCTGCAATGGATGAATGAAGC 1655
DB 1 ACTGGAACTGGCAGAGAGGCTCTGGCTTCCAAACANNNNNNNNNNNNNNNNNNNNN 60

QY 1656 AAACCATTCGCAAGCAGAGAGGACCTGGAACACCATGACCAT-CCTCAGGGCTCAGATG 1714
DB 61 NNNNNNNNNNGCAGGAAGAGGACCTGGAACCATGACCATCNNNNNNNNNNNNNNN 120

QY 1715 GAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCAGCAGAGAAATTCATGAGGAA 1774
DB 121 NNN 180

QY 1775 AAGGAGCACTGCAATTCGAGCTGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTCAAGAC 1834
DB 181 AAGGAGCACTGNN 240

QY 1835 GAGGAGCAGGAGCTCTTGTATGAGATGACAGAGTGGTATGGGGCGAGAACAACTGACTCT 1894
DB 241 NNAGGCAGGCAGTCTTGTATGAGATGACAGAGTGGTATGGGGCGAGAACAACTGACTCT 300

QY 1895 GACAGCAGGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGAGCTGGGGCGAACAGCGG 1954
DB 301 GACAGCAGGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGAGCTGGGGCGAACAGCGG 360

QY 1955 AATATTCGATTCATTCCTGCCCCCAAGTGAGAGGTTCTGCTGACATAGACAGTTA 2014
DB 361 AATATTCGATTCATTCCTGCCCCCAAGTGAGAGGTTCTGCTGACATAGACAGTTA 420

QY 2015 CAGATTCACGTGATGATTCATCAATTAAGTGTGTATGATGATCACTCCCAAACTGTT 2074
DB 421 CAGATTCACGTGATGATTCATCAATTAAGTGTGTATGATGATCACTCCCAAACTGTT 480

QY 2075 GGT 2077
DB 481 GGT 483

RESULT 44
ACH20402
ID ACH20402 standard; cDNA; 432 BP.
AC ACH20402;
XX
XX
DT 13-OCT-2003 (first entry)
XX
XX Human adult liver cDNA #14.
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.

PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 7614; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 432 BP; 120 A; 129 C; 121 G; 62 T; 0 U; 0 Other;

Query Match 14.3%; Score 296.2; DB 9; Length 432;
Best Local Similarity 99.0%; Pred. No. 1.1e-69;
Matches 298; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 298 AGGAACCTTCTGCAATGTCCTCATCAACCTCTCAGCTGCCTCACTGAAAGAGGACAGCCC 357
DB 130 AGGAACCTTCTGCAATGTCCTCATCAACCTCTCAGCTGCCTCACTGAAAGAGGACAGCCC 189

QY 358 CAGTGAAGCAGCAGGAATGGACCCCTCCCTGGCCCAACCTGACAGAGCTTTAC 417
DB 190 CAGTGAAGCAGCAGGAATGGACCCCTCCCTGGCCCAACCTGACAGAGCTTTAC 249

QY 418 CCGGAGAGCTGCTGCAGCAGATGAAGAGCTCTCTGAACCAAGAACCCAGCTGAAAGA 477
DB 250 CCGGAGAGCTGCTGCAGCAGATGAAGAGCTCTCTGACCGAGAACCCAGCTGAAAGA 309

QY 478 AGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTCGGCCTGGAC 537
DB 310 AGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTCGGCCTGGAC 369

QY 538 AGAACAACAGAGGAAGAACGCCAGTTTGTGATACAGAGCAAGAAAGAGAGCG 597
DB 370 AGAACAACAGAGGAAGAACGCCAGTTTGTGATACAGAGCAAGAAATGAGCG 429

QY 598 T 598
DB 430 T 430

RESULT 45
AEE13195
ID AEE13195 standard; DNA; 505 BP.
XX
XX AEE13195;
AC AEE13195;
XX
XX 26-JAN-2006 (first entry)
XX
XX Hamster consensus sequence SEQ ID NO 2706.
XX

```
KW array: DNA sequencing; ds.
XX Cricetulus griseus.
XX WO2005111246-A1.
XX 24-NOV-2005.
XX 11-MAY-2005; 2005WO-US016880.
XX 11-MAY-2004; 2004US-0570425P.
XX (AMHP ) WYETH.
XX Melville MW, Charlebois TS, Mounts WM, Hann LE, Sinacore MS;
XX Leonard MW, Brown EL, Miller CP, Lee GW;
XX WPI; 2005-811678/82.
XX
XX Forming an oligonucleotide array directed toward an unsequenced organism
XX comprises selecting oligonucleotide probes comprising a first set of
XX oligonucleotide probes, each of which is specific for one of the template
XX sequences.
XX
XX Claim 5; SEQ ID NO 2706; 68pp; English.
XX
XX The invention relates to a method of forming an oligonucleotide array
XX directed toward an unsequenced organism comprises selecting
XX oligonucleotide probes comprising a first set of oligonucleotide probes,
XX each of which is specific for one of the template sequences. The method
XX is useful for forming an oligonucleotide array directed toward an
XX unsequenced organism. The array is useful for detecting the presence,
XX absence, and/or quantity of expression levels of genes in a cell derived
XX from an unsequenced organism. The present sequence represents a consensus
XX sequence for undiscovered CHO cell genes.
XX
XX Sequence 505 BP; 155 A; 122 C; 141 G; 87 T; 0 U; 0 Other;
XX
Query Match 14.1%; Score 292.4; DB 14; Length 505;
Best Local Similarity 80.8%; Pred. No. 1.3e-68;
Matches 341; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 289 CTTTTCACAGGAACCTTGTGCAATGTCCCATCAACTCTCAGTGTGCTACTGAAAGGA 348
DB 83 CTAAGCCGAGCGACCTGTACCATGTCCCATCAACTCTCGAGCTGCTGACTGAGAAGGG 142
QY 349 GGACAGCCCCAGTGAAGCACAGGAATGGACCCCTCCACCTGGCCACCCAAACCTTGA 408
DB 143 TGACAGCCTCTGTGAGACCCCGAGGAATGAACCCCTTAATAAGGTTTCAACCCAGCTTGA 202
QY 409 CAGCTTTACCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAACCA 468
DB 203 CACATTCACCTCTGAGGAGATGCTGCAGCAATGAAGGAATCTCTGTTGAGAACCA 262
QY 469 GCTGAAAGAGCCATGAAGCTAAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTC 528
DB 263 GCTGAAAGAGCCATGAAGCTAAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTC 322
QY 529 GGCCTGGACAGAGAAACAGAGGAAGCAAGCCAGTTTGTGATATCAGAGCAAGCAAGC 588
DB 323 AGCCTGGCAGAGAGCAGAGGAAGCAAGCAAGCTGTTGTTGACACACAGCAAGCAAGC 382
QY 589 AAGAGCGCTCTAATGGCTTGTAGTCATGAGATGAGAAATTCAGAGAGCTTGGAAA 648
DB 383 TAAGAAGCTCTGATGACCTTGTAGTTATGAAATGAGAACTGAAGGTGAGCTTGGAAA 442
QY 649 ACTAAAGGGAATCAGAAAGGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCCGAG 708
DB 443 ACTAAGAGAGAGTCAAGAAAGGCCATTTGAAGGCTCCATATGTGGCTCCAGGCTCCCAT 502
QY 709 GG 710
DB 503 GG 504
```

```
RESULT 46
AEE16837
ID AEE16837 standard; DNA; 505 BP.
XX
XX AEE16837;
XX
XX 26-JAN-2006 (first entry)
XX
XX Hamster SEQ 2706 tiling sequence SEQ ID NO 6348.
XX
XX array; DNA sequencing; ds.
XX
XX Cricetulus griseus.
XX
XX WO2005111246-A1.
XX
XX 24-NOV-2005.
XX
XX 11-MAY-2005; 2005WO-US016880.
XX
XX 11-MAY-2004; 2004US-0570425P.
XX
XX (AMHP ) WYETH.
XX
XX Melville MW, Charlebois TS, Mounts WM, Hann LE, Sinacore MS;
XX Leonard MW, Brown EL, Miller CP, Lee GW;
XX WPI; 2005-811678/82.
XX
XX Forming an oligonucleotide array directed toward an unsequenced organism
XX comprises selecting oligonucleotide probes comprising a first set of
XX oligonucleotide probes, each of which is specific for one of the template
XX sequences.
XX
XX Claim 5; SEQ ID NO 6348; 68pp; English.
XX
XX The invention relates to a method of forming an oligonucleotide array
XX directed toward an unsequenced organism comprises selecting
XX oligonucleotide probes comprising a first set of oligonucleotide probes,
XX each of which is specific for one of the template sequences. The method
XX is useful for forming an oligonucleotide array directed toward an
XX unsequenced organism. The array is useful for detecting the presence,
XX absence, and/or quantity of expression levels of genes in a cell derived
XX from an unsequenced organism. The present sequence represents a tiling
XX sequence of one of the sequences disclosed in the specification used to
XX generate the probes of the invention.
XX
XX Sequence 505 BP; 155 A; 122 C; 141 G; 87 T; 0 U; 0 Other;
XX
Query Match 14.1%; Score 292.4; DB 14; Length 505;
Best Local Similarity 80.8%; Pred. No. 1.3e-68;
Matches 341; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 289 CTTTTCACAGGAACCTTGTGCAATGTCCCATCAACTCTCAGTGTGCTACTGAAAGGA 348
DB 83 CTAAGCCGAGCGACCTGTACCATGTCCCATCAACTCTCGAGCTGCTGACTGAGAAGGG 142
QY 349 GGACAGCCCCAGTGAAGCACAGGAATGGACCCCTCCACCTGGCCACCCAAACCTTGA 408
DB 143 TGACAGCCTCTGTGAGACCCCGAGGAATGAACCCCTTAATAAGGTTTCAACCCAGCTTGA 202
QY 409 CAGCTTTACCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAACCA 468
DB 203 CACATTCACCTCTGAGGAGATGCTGCAGCAATGAAGGAATCTCTGTTGAGAACCA 262
QY 469 GCTGAAAGAGCCATGAAGCTAAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTC 528
DB 263 GCTGAAAGAGCCATGAAGCTAAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTC 322
QY 529 GGCCTGGACAGAGAAACAGAGGAAGCAAGCCAGTTTGTGATATCAGAGCAAGCAAGC 588
```


FT /tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT replace(894,T)
FT /tag= e
FT /standard_name= "Single nucleotide polymorphism"
FT replace(987,C)
FT /tag= f
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1112,C)
FT /tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1505,CC)
FT /tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1606,A)
FT /tag= i
FT /standard_name= "Single nucleotide polymorphism"
FT replace(2405,T)
FT /tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT replace(2606,G)
FT /tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT replace(3313,A)
FT /tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT replace(3555,TT)
FT /tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT replace(3625,G)
FT /tag= n
FT /standard_name= "Single nucleotide polymorphism"
FT replace(3629,C)
FT /tag= o
FT /standard_name= "Single nucleotide polymorphism"
FT replace(3882,TT)
FT /tag= p
FT /standard_name= "Single nucleotide polymorphism"
FT replace(3988,T)
FT /tag= q
FT /standard_name= "Single nucleotide polymorphism"
FT replace(4452,A)
FT /tag= r
FT /standard_name= "Single nucleotide polymorphism"
XX
XX US2003190617-A1.
XX
XX 09-OCT-2003.
XX
XX 06-MAR-2002; 2002US-00091281.
XX
XX 06-MAR-2002; 2002US-00091281.
XX
XX (SIEB/) SI E.
XX (RAYM/) RAYMOND V.
XX (MORI/) MORISSETTE J.
XX
XX Raymond V, Morissette J, Si E;
XX WPI; 2003-864168/80.
XX
XX New nucleic acid sequences of the optineurin gene are useful to detect
XX polymorphisms particularly single nucleotide polymorphisms in the
XX optineurin promoter to diagnose, prognosis and treat glaucoma and related
XX disorders.
XX
XX Disclosure; SEQ ID NO 2; 159pp; English.
XX
XX The invention relates to an isolated nucleic acid (N1) comprising at
XX least 20 but not more than 1500 consecutive nucleotides of the optineurin
XX promoter appearing as AD813890. Also included are the optineurin promoter
XX operably linked to a heterologous nucleic acid, a nucleic acid capable of
XX detecting a single nucleotide polymorphism (SNP) in the optineurin

CC promoter, a host cell comprising the promoter operably linked to a
CC heterologous sequence, diagnosing or prognosing glaucoma in a sample
CC obtained from a cell or bodily fluid (comprising detecting a polymorphism
CC in a promoter region of the optineurin gene associated with a glaucoma
CC phenotype), detecting a SNP sequence variation in a sample containing
CC DNA, detecting the presence of an optineurin promoter sequence variation
CC in a sample containing DNA, determining the presence or increased
CC susceptibility to glaucoma or to a progressive ocular hypertensive
CC disorder resulting in loss of visual field in a patient (or the severity
CC or progression of glaucoma in a patient, comprising providing
CC amplification reaction primers that direct amplification of a selected
CC nucleic acid region containing the variation within the optineurin
CC promoter and amplifying the DNA) and detecting a polymorphism (comprising
CC obtaining a sample containing human genomic DNA, providing a nucleic acid
CC capable of detecting a SNP located within an optineurin promoter, and
CC detecting the polymorphism). The invention is used to diagnose and
CC prognose glaucoma and also to treat glaucoma related disorders. The
CC present sequence is the optineurin gene.
XX
SQ Sequence 46951 BP; 12703 A; 10108 C; 10051 G; 14089 T; 0 U; 0 Other;
Query Match 9.9%; Score 206; DB 10; Length 46951;
Best Local Similarity 100.0%; Pred. No. 4.2e-44; Mismatches 0; Gaps 0;
Matches 206; Conservative 0; Indels 0;
QY 475 AGAAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTGCGCCTG 534
Db 15116 AGAAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTGCGCCTG 15175
QY 535 GACAGAGAAACAGAGGAAGAACGCCAGTTTTTTTGTAGATACAGAGCAAGAGCAAAAGA 594
Db 15176 GACAGAGAAACAGAGGAAGAACGCCAGTTTTTTTGTAGATACAGAGCAAGAGCAAAAGA 15235
QY 595 GCGTCTAATGGCTTTGAGTCTATGAGATGAGAAATTTGAAGGAGAGCTTTGGAATACTAAA 654
Db 15236 GCGTCTAATGGCTTTGAGTCTATGAGATGAGAAATTTGAAGGAGAGCTTTGGAATACTAAA 15295
QY 655 AGGGAATCAGAAAGTCTATCTGAGG 680
Db 15296 AGGGAATCAGAAAGTCTATCTGAGG 15321

Search completed: May 30, 2006, 01:44:33
Job time : 2353 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2006, 23:56:43 ; Search time 8892 Seconds
(without alignments)

13061.679 Million cell updates/sec

Title: SEQ1-458A

Perfect score: 2077

Sequence: 1 atcccggtcggtggattctctt.....acctcccaaaactgttggt 2077

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : EST:*

1: gb_est1:*

2: gb_est3:*

3: gb_est4:*

4: gb_est5:*

5: gb_est6:*

6: gb_estc:*

7: gb_est2:*

8: gb_est7:*

9: gb_est8:*

10: gb_est9:*

11: gb_gss1:*

12: gb_gss2:*

13: gb_gss3:*

14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1770.8	85.3	1864	6	CR601600
2	1738.8	83.7	1862	6	CR597026
3	1710.8	82.4	1846	6	CR603260
4	1426.2	68.7	3265	6	CR859573
5	1240.4	59.7	1612	14	DQ036021
6	986.6	47.5	1089	4	BX384735
7	982.6	47.3	1068	1	AL583163
8	957.6	46.1	1032	1	AL570028
9	937.2	45.1	1021	1	AL562664
10	902.2	43.4	1004	1	AL533761
11	882.4	42.5	1012	4	BX343673
12	868.8	41.8	996	1	AL516921
13	857.6	41.3	984	1	AL526184
14	837.6	40.3	1001	1	AL583164
15	824.4	39.7	1497	14	DQ036022
16	806.8	38.8	810	9	CX781622
17	800	38.5	882	1	AL565248
18	797	38.4	955	1	AL516922
19	789.2	38.0	882	3	BU161830

20	777.2	37.4	968	1	AL543413
21	772	37.2	772	8	CN309549
22	771.4	37.1	817	3	BU623283
23	770	37.1	945	7	BE875862
24	763.8	36.8	854	5	CD243803
25	747.6	36.0	990	4	BX364738
26	746.2	35.9	919	1	AL565831
27	744.8	35.9	900	4	BX439653
28	730	35.1	761	5	CF456387
29	728	35.1	989	1	AL538121
30	725.6	34.9	872	3	BO883124
31	721.4	34.7	907	1	AL535231
32	717.6	34.5	1201	10	DV778070
33	715	34.4	921	3	BQ951546
34	704.4	33.9	975	3	BQ650013
35	697.4	33.6	878	3	BQ218575
36	694	33.4	696	8	CN309548
37	693	33.4	733	3	BU630531
38	690.8	33.3	904	3	BQ879423
39	686	33.0	985	1	AL567151
40	682	32.8	723	3	BQ773529
41	680.8	32.8	726	3	BQ776232
42	680.4	32.8	742	3	BU628287
43	679.8	32.7	871	5	CD245471
44	670.8	32.3	740	2	BG031732
45	664.2	32.0	895	1	AU122473
46	659.8	31.8	664	2	BI062019
47	659.8	31.8	891	9	DN110950
48	649	31.2	875	9	DN111300
49	643	31.0	693	3	BU626293
50	640.4	30.8	905	3	BQ893002
51	639	30.8	639	3	BM931090
52	639	30.8	639	8	CA482005
53	635	30.6	674	5	CD742649
54	628.2	30.2	1200	6	AK015354
55	624.4	30.1	627	8	CV030612
56	623	30.0	857	8	CV979174
57	619.4	29.8	681	5	CK003305
58	619.4	29.8	707	8	CN309550
59	617.8	29.7	954	5	CD557312
60	616.4	29.7	619	8	CV027807
61	608.8	29.3	826	5	CK848814
62	608	29.3	608	1	AL134550
63	605.4	29.1	659	2	BG119942
64	603.2	29.0	735	5	CD356996
65	601	28.9	2327	6	BC024755
66	599.4	28.9	640	4	CA416446
67	596.4	28.7	598	3	BP309106
68	595.2	28.7	841	3	BP437034
69	584.4	28.1	597	8	CV570003
70	582.2	28.0	800	7	BE884512
71	582	28.0	582	3	BP207397
72	581.8	28.0	651	9	CX781554
73	581.4	28.0	583	4	CB111029
74	580.4	28.0	583	9	DA601485
75	580.4	27.9	641	3	BU625819
76	579.2	27.9	588	3	BP249758
77	579	27.9	579	9	DA168117
78	578	27.8	852	8	CR787930
79	576	27.7	587	3	BP218417
80	575.8	27.7	642	1	AA196381
81	573.4	27.6	603	4	CB141129
82	572.2	27.5	577	3	BP263098
83	568.2	27.4	822	5	CJ431285
84	566.4	27.3	651	2	BG926123
85	565.8	27.2	585	3	BP205716
86	563.8	27.1	582	3	BP215313
87	561.2	27.0	589	4	CB265751
88	561	27.0	573	8	CN309553
89	558.2	26.9	718	2	BI828447
90	557.8	26.9	682	7	BE729841
91	557.4	26.8	567	3	BP204471
92	557.4	26.8	581	3	BP260233

AL543413	AL543413
CN309549	170005998
BU623283	UI-H-FL1-
BE875862	601487327
CD243803	AGENCOURT
BX364738	BX364738
AL565831	AL565831
AX439653	AX439653
CF456387	AGENCOURT
AL538121	AL538121
BO883124	AGENCOURT
AL535231	AL535231
DV778070	Hw Fat 56
BQ951546	AGENCOURT
BQ650013	AGENCOURT
BQ218575	AGENCOURT
CN309548	170005829
BU630531	UI-H-FL0-
AL567151	AL567151
BO773529	UI-H-FE0-
BO776232	UI-H-FE0-
BU628287	UI-H-FG0-
CD245471	AGENCOURT
BG031732	602300889
AU122473	AU122473
BI062019	IL3-UT011
DN110950	1110027 M
DN111300	1110411 M
BU626293	UI-H-DF0-
BQ893002	AGENCOURT
BM931090	UI-E-EJ0-
CA482005	hw14h10.Y
CD742649	UI-H-FI2-
AK015354	Mus muscu
CV030612	9787 Full
CV979174	UMC-bemiv
CK003305	AGENCOURT
CN309550	170004710
CD557312	AGENCOURT
CV027807	6157 Full
CK848814	971625 BA
AL134550	DKF2P547P
BG119942	602352210
CD356996	AGENCOURT
BC024755	Mus muscu
CA416446	UI-H-FE0-
BP309106	BP309106
BP437034	BP437034
CV570003	od12a10.Y
BE884512	601506208
BP207397	BP207397
CX781554	HESC3_7.D
CB111029	K-EST0152
DA601485	DA601485
BU625819	UI-H-FG1-
BP249758	BP249758
DA168117	DA168117
CR787930	DKF2P459C
BP218417	BP218417
AA196381	2p99g11.8
CB141129	K-EST0194
BP263098	BP263098
CJ431285	CJ431285
BG926123	HNC32-1-F
BP205716	BP205716
BP215313	BP215313
CB265751	1004656 H
CN309553	170004241
BI828447	603078213
BE729841	601564758
BP204471	BP204471
BP260233	BP260233


```
Db 631 ATTCCTTTGTTGAAATTTAGGATGCGTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA 690
Qy 897 AGCATAGTCTGGGCCCAACGAGAAACAGTCTCCACTGGCAGCGCAATTTCTAAATATAGGA 956
Db 691 AGCATAGTCTGGGCCCAACGAGAAACAGTCTCCACTGGCAGCGCAATTTCTAAATATAGGA 750
Qy 957 GCAGATCTGCAGATGGGGCCAGAAATTTACTTCGAAACATGAGGATTAACCTGTGAGCCAGC 1016
Db 751 GCAGATCTGCAGATGGGGCCAGAAATTTACTTCGAAACATGAGGATTAACCTGTGAGCCAGC 810
Qy 1017 TCCTGCTGTGCTAAAGGGAAGGGAATCAGAAGGTGGAGAGACTTTGAAGTTGCACTCAAGG 1076
Db 811 TCCTGCTGTGCTAAAGGGAAGGGAATCAGAAGGTGGAGAGACTTTGAAGTTGCACTCAAGG 870
Qy 1077 AGGCCAAAGAAAGATTTTCAGATTTTCAAAGAAACAAAGTAATCGTTCTGAGATTGAAA 1136
Db 871 AGGCCAAAGAAAGATTTTCAGATTTTCAAAGAAACAAAGTAATCGTTCTGAGATTGAAA 930
Qy 1137 CCAGACAGAGGGGGACACAGAGAAAGAGAAATGATGAAGAAAGGCCCGGAGACTGTTG 1196
Db 931 CCAGACAGAGGGGGACACAGAGAAAGAGAAATGATGAAGAAAGGCCCGGAGACTGTTG 990
Qy 1197 GAAGCGAAGTGGGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGG 1256
Db 991 GAAGCGAAGTGGGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGG 1050
Qy 1257 CTCATACAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTGAGG 1316
Db 1051 CTCATACAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTGAGG 1110
Qy 1317 CCCTTGAAGGAAAAATTTGCAATTTCCATCAGAGTTGAATGAAGAAAGAGCTGTTT 1376
Db 1111 CCCTTGAAGGAAAAATTTGCAATTTCCATCAGAGTTGAATGAAGAAAGAGCTGTTT 1170
Qy 1377 ATACTAAGAAAAAGTTAGCTCAAGTGAAGAGCTGCTATCAGAAATCAAAATGAAC 1436
Db 1171 ATACTAAGAAAAAGTTAGCTCAAGTGAAGAGCTGCTATCAGAAATCAAAATGAAC 1230
Qy 1437 AGGCTAAACAGAGGATGAAAGTCCAAATTAACCTGTGCTACAGATGACACACAAAGC 1496
Db 1231 AGGCTAAACAGAGGATGAAAGTCCAAATTAACCTGTGCTACAGATGACACACAAAGC 1290
Qy 1497 TTCTTCAAGAACATATATGCAATTTGAAACCAATTTGAGGAACCTAACAGAAAAAGAGTCAG 1556
Db 1291 TTCTTCAAGAACATATATGCAATTTGAAACCAATTTGAGGAACCTAACAGAAAAAGAGTCAG 1350
Qy 1557 AAAAGTGGACAGGCGAGTCTGAGGAACCTGAGTGAAGAACTGGAACTGGCAGAGAGG 1616
Db 1351 AAAAGTGGACAGGCGAGTCTGAGGAACCTGAGTGAAGAACTGGAACTGGCAGAGAGG 1410
Qy 1617 CTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAAG 1676
Db 1411 CTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAAG 1470
Qy 1677 AGGACCTGGAAAACCATGACCATCTCAGGCGCTCAGATGGAAGTTTACTGTTCTGATTTTC 1736
Db 1471 AGGACCTGGAAAACCATGACCATCTCAGGCGCTCAGATGGAAGTTTACTGTTCTGATTTTC 1530
Qy 1737 ATGCTGAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGAGGAACTGGCAATGGAGC 1796
Db 1531 ATGCTGAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGAGGAACTGGCAATGGAGC 1590
Qy 1797 TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTCAAGACGGAGGCGAGCTCTCTGATGG 1856
Db 1591 TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTCAAGACGGAGGCGAGCTCTCTGATGG 1650
Qy 1857 AGATGAGAGTCTGATGGGGCGAGAAACAAAGTGAATCTGACACGAGCGGCTTACTTGTGTC 1916
Db 1651 AGATGAGAGTCTGATGGGGCGAGAAACAAAGTGAATCTGACACGAGCGGCTTACTTGTGTC 1710
Qy 1917 AAAGAGGAGCTGAGGACAGGGAATGCGGGCAACAGCGGGAATATTCGATTCATTCCTGCC 1976
Db 1711 AAAGAGGAGCTGAGGACAGGGAATGCGGGCAACAGCGGGAATATTCGATTCATTCCTGCC 1770
```

```
Qy 1977 CCAAGTGTGAGAGGTTCTGCTGACATAGACAGTTTACAGATTCACTGATGATTGCA 2036
Db 1771 CCAAGTGTGAGAGGTTCTGCTGACATAGACAGTTTACAGATTCACTGATGATTGCA 1830
Qy 2037 TCATTTAAGTGTGATGATATCACCTCCCAAAAC 2070
Db 1831 TCATTTAAGTGTGATGATATCACCTCCCAAAAC 1864

RESULT 2
CR597026 1862 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DL009YF13 of B cells (Ramos cell line)
DEFINITION Cot 25-normalized of Homo sapiens (human).
ACCESSION CR597026
VERSION 1 GI:50477833
KEYWORDS HTC; CNSLT_cDNA..
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1862)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1862)
Genoscope.
REFERENCE 2 (bases 1 to 1862)
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 131 91066 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1..1862
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YF13"
/tissue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 83.7%; Score 1738.8; DB 6; Length 1862;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 297 CAGGAATCTCTGCAATGTCCCACTCAACCTCTCAGCTGCCTCACTGAAAAGGAGACAGCC 356
Db 121 CAGGAATCTCTGCAATGTCCCACTCAACCTCTCAGCTGCCTCACTGAAAAGGAGACAGCC 180
Qy 357 CCAGTGAAGACAGAGAAATGGACCCCTCCACCTGGCCCAACCCCAACCTGGACACGTTTA 416
Db 181 CCAGTGAAGACAGAGAAATGGACCCCTCCACCTGGCCCAACCCCAACCTGGACACGTTTA 240
Qy 417 CCCCGAGGAGCTGTGCGAGCAGATGAAAGAGCTCTGACCAAGAACCCACCTGTAAG 476
Db 241 CCCCGAGGAGCTGTGCGAGCAGATGAAAGAGCTCTGACCAAGAACCCACCTGTAAG 300
Qy 477 AAGCCATGAGCTTAATATCAAGCCATGAAGGGAGATTGAGGAGCTTTCGGCTTGA 536
Db 301 AAGCCATGAGCTTAATATCAAGCCATGAAGGGAGATTGAGGAGCTTTCGGCTTGA 360
Qy 537 CAGAGAAAACAGAGGAGAAAGCCAGCTTTTTCAGATACAGAGCAAGAAAGCAAGAGAGC 596
```

Db	361	CAGAGAAACAGAGGAAGAACCGCAGTTTTTTTGTAGATACAGAGCAAAAGAACGAAAGAGC	420
QY	597	GTCTAATGCGCTTTGACTCATGAGATGAGAAATTGAAGGAAGAGCTTTGGAAGAACTAAAG	656
Db	421	GTCTAATGCGCTTTGACTCATGAGATGAGAAATTGAAGGAAGAGCTTTGGAAGAACTAAAG	480
QY	657	GGAAATCAGAAAGGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAG	716
Db	481	GGAAATCAGAAAGGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAG	540
QY	717	CGGAGCAGAAAGGACGAGCTCAGAACCCAGGTGTGAGGCTTCAAGCAGAGAGGCGAG	776
Db	541	CGGAGCAGAAAGGACGAGCTCAGAACCCAGGTGTGAGGCTTCAAGCAGAGAGGCGAG	600
QY	777	ACCTGTTGGGCATCGTGTCTGAACCTCAGCTCAAGCTGAACCTCCAGCGGCTCCTCAGAG	836
Db	601	ACCTGTTGGGCATCGTGTCTGAACCTCAGCTCAAGCTGAACCTCCAGCGGCTCCTCAGAG	660
QY	837	ATTCTCTTCTGTAATTAGGATGGCTGAAAGGAGAACGAGAAAGGTCAGTAAAGAAATCA	896
Db	661	ATTCTCTTCTGTAATTAGGATGGCTGAAAGGAGAACGAGAAAGGTCAGTAAAGAAATCA	720
QY	897	AGCATAGTCCTGGGCCACGAGAACAGTCTCCACTGGCACGGCATTTGTCTAAATATAGGA	956
Db	721	AGCATAGTCCTGGGCCACGAGAACAGTCTCCACTGGCACGGCATTTGTCTAAATATAGGA	780
QY	957	GCAGATCTCAGATGGGCCCAAGAAATTAATTCGAACTAGGAGTTAACTGTGAGCCAGC	1016
Db	781	GCAGATCTCAGATGGGCCCAAGAAATTAATTCGAACTAGGAGTTAACTGTGAGCCAGC	840
QY	1017	TCTGCTGTGCTTAAGGGAAGGAATCAGAAAGTGAGAGACTTGAAGTTGCACTCAAGG	1076
Db	841	TCTGCTGTGCTTAAGGGAAGGAATCAGAAAGTGAGAGACTTGAAGTTGCACTCAAGG	900
QY	1077	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCTCTTGAGATTGAAA	1136
Db	901	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCTCTTGAGATTGAAA	960
QY	1137	CCGAGCAGAGGGGACACAGAGAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTG	1196
Db	961	CCGAGCAGAGGGGACACAGAGAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTG	1020
QY	1197	GAGCGAAGTGGAGACACTGAACCTCCAGGTGACATCTCTTTTAAGGAGCTTCAAGAGG	1256
Db	1021	GAGCGAAGTGGAGACACTGAACCTCCAGGTGACATCTCTTTTAAGGAGCTTCAAGAGG	1080
QY	1257	CTCATACAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTFCAGG	1316
Db	1081	CTCATACAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTFCAGG	1140
QY	1317	CCCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAGCAAGAGCTTGT	1376
Db	1141	CCCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAGCAAGAGCTTGT	1200
QY	1377	ATACTAAACAAAGTTAGAGCTACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAAC	1436
Db	1201	ATACTAAACAAAGTTAGAGCTACAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAAC	1260
QY	1437	AGGCTAAACAGAGGATGAAAGTCCAAATTAATCTGTCTACAGATGACACACAAGC	1496
Db	1261	AGGCTAAACAGAGGATGAAAGTCCAAATTAATCTGTCTACAGATGACACACAAGC	1320
QY	1497	TTCTTCAAGAACATAATATGCAATTAAGAACATTTGAGCACTAAACAGAAAGAGTCAG	1556
Db	1321	TTCTTCAAGAACATAATATGCAATTAAGAACATTTGAGCACTAAACAGAAAGAGTCAG	1380
QY	1557	AAAAAGTGGACAGGGCAGTGTCTGAAGAACTGAGTGAAGAACTGGAACTGGCAGAGAGG	1616
Db	1381	AAAAAGTGGACAGGGCAGTGTCTGAAGAACTGAGTGAAGAACTGGAACTGGCAGAGAGG	1440
QY	1617	CTCTGCTTCCAAACAGCTGCAAAATGGAATGAAGCAAAACATTTGCCAAGCAGGAAG	1676
Db	1441	CTCTGCTTCCAAACAGCTGCAAAATGGAATGAAGCAAAACATTTGCCAAGCAGGAAG	1500
QY	1677	AGGACTTGAAACACCATGACCATCTCAGGGCTCAGATGGAAGCTTACTGTCTGATTTTC	1736
Db	1501	AGGACTTGAAACACCATGACCATCTCAGGGCTCAGATGGAAGCTTACTGTCTGATTTTC	1560
QY	1737	ATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGAGCAACTTGGCATTCGAGC	1796
Db	1561	ATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGAGAAAAGAGCAACTTGGCATTCGAGC	1620
QY	1797	TGGCAGTTCTGCTGAAAGAGATGATGCTTTCGAGAGCGGAGCGGAGCTCCTTCATGG	1856
Db	1621	TGGCAGTTCTGCTGAAAGAGATGATGCTTTCGAGAGCGGAGCGGAGCTCCTTCATGG	1680
QY	1857	AGATCAGAGTCTCATGGGGCGAGAAACAAAGTGAATCTCACCAGCAGGCTTACCTTGTTC	1916
Db	1681	AGATCAGAGTCTCATGGGGCGAGAAACAAAGTGAATCTCACCAGCAGGCTTACCTTGTTC	1740
QY	1917	AAAGAGGAGCTGAGGACAGGGACTCGCGGCAACAGCGGAATATTCGATTCATTCCTGCC	1976
Db	1741	AAAGAGGAGCTGAGGACAGGGACTCGCGGCAACAGCGGAATATTCGATTCATTCCTGCC	1800
QY	1977	CAAAGTGGGAGAGCTTCTGCTGACATAGACACGTTACAGATTCACGTCATGGATTGCA	2036
Db	1801	CAAAGTGGGAGAGCTTCTGCTGACATAGACACGTTACAGATTCACGTCATGGATTGCA	1860
QY	2037	TC 2038	
Db	1861	TC 1862	
RESULT 3			
CR603260			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local similarity			

Matches 1740; Conservative 0; Mismatches 2; Indels 18; Gaps 1;			
Qy	297	CAGGAACTTCTGCAATGCCATCAACCTCTCAGCTGCTCCTCACTGAAAGGAGGACGACC	356
Db	105	CAGGAACTTCTGCAATGCCATCAACCTCTCAGCTGCTCCTCACTGAAAGGAGGACGACC	164
Qy	357	CCAGTGAAGCAGAGAAATGGACCCGCCCCACCTTGGGCCACCCCAAACTTGACACAGTTT	416
Db	165	CCAGTGAAGCAGAGAAATGGACCCGCCCCACCTTGGGCCACCCCAAACTTGACACAGTTT	224
Qy	417	CCCCGAGGAGCTCTGCGACGATGAAAGAGCTCTGACCAAGAACCCACCAGCTGAAAG	476
Db	225	CCCCGAGGAGCTCTGCGACGATGAAAGAGCTCTGACCGGAACACACAGCTGAAAG	284
Qy	477	AAGCCATGAAGCTAAATTAATCAAGCCATGAAAGGAGATTTGAGGAGCTTTCCGCCCTGGA	536
Db	285	AAGCCATGAAGCTAAATTAATCAAGCCATGAAAGGAGATTTGAGGAGCTTTCCGCCCTGGA	344
Qy	537	CAGAGAACAGAGGAGGAGGAGCCAGCTTTTGTGATACAGAGCAAAAGAGCAAAAGAGC	596
Db	345	CAGAGAACAGAGGAGGAGGAGGAGCCAGCTTTTGTGATACAGAGCAAAAGAGCAAAAGAGC	404
Qy	597	GTCTAATGGCTTGAGTCATGAGAAATGAGAAATTTGAAGGAGAGCTTTGGAAGCTTAAAG	656
Db	405	GTCTAATGGCTTGAGTCATGAGAAATGAGAAATTTGAAGGAGAGCTTTGGAAGCTTAAAG	464
Qy	657	GGAAATCAGAAAGGTCTCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAAG	716
Db	465	GGAAATCAGAAAGGTCTCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAAG	524
Qy	717	CGAGCAGGAAAGGACAGCTCAGGACCCAGAGTGTGAGGCTTCAAGCAGAGAGGCGAG	776
Db	525	CGAGCAGGAAAGGACAGCTCAGGACCCAGAGTGTGAGGCTTCAAGCAGAGAGGCGAG	584
Qy	777	ACCTGTTGGCATCTGCTGTAACCTGAGCTCAAGCTGAACTCCAGGGCTCCTCAGAAG	836
Db	585	ACCTGTTGGCATCTGCTGTAACCTGAGCTCAAGCTGAACTCCAGGGCTCCTCAGAAG	644
Qy	837	ATTCTCTTTGTAATTAAGATGGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA	896
Db	645	ATTCTCTTTGTAATTAAGATGGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA	704
Qy	897	AGCATAGTCTGGGCCCAAGAGAACTGCTCCTGAGCAGGCAATGCTTAAATATAGGA	956
Db	705	AGCATAGTCTGGGCCCAAGAGAACTGCTCCTGAGCAGGCAATGCTTAAATATAGGA	746
Qy	957	GCAGATCTGAGATGGGGCCCAAGAAATTTCTCGAACAATGAGGATTAACCTGTAGCCAGC	1016
Db	747	GCAGATCTGAGATGGGGCCCAAGAAATTTCTCGAACAATGAGGATTAACCTGTAGCCAGC	806
Qy	1017	TCCTGCTGTCTTAAGGAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGG	1076
Db	807	TCCTGCTGTCTTAAGGAGGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCAAGG	866
Qy	1077	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGGAAACCAAGTAATCGTTCTGAGATGAAA	1136
Db	867	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGGAAACCAAGTAATCGTTCTGAGATGAAA	926
Qy	1137	CCGAGCAGAGGGGACACAGAGAGAGAAATGATGAGAGAAAGGCCCGGAGACTGTG	1196
Db	927	CCGAGCAGAGGGGACACAGAGAGAGAAATGATGAGAGAAAGGCCCGGAGACTGTG	986
Qy	1197	GAAGCGAAGTGGAGACACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGG	1256
Db	987	GAAGCGAAGTGGAGACACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGG	1046
Qy	1257	CTCATCAAAACTCAGCGAAAGCTGAGCTAATGAAGAGAGACCTTCAAGAAAGTGTCCAGG	1316
Db	1047	CTCATCAAAACTCAGCGAAAGCTGAGCTAATGAAGAGAGACCTTCAAGAAAGTGTCCAGG	1106
Qy	1317	CCCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAGCAGAGAGCTTGT	1376
Db	1107	CCCTTGAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAGCAGAGAGCTTGT	1166

Qy	1377	ATACTAACAACAAAGTTAGAGCTACAAGTGGAAACATGCTATCAGAATCAAAATGGAAC	1436
Db	1167	ATACTAACAACAAAGTTAGAGCTACAAGTGGAAACATGCTATCAGAATCAAAATGGAAC	1226
Qy	1437	AGGCTAAAAACAGAGGATGAAAGATCCAAATTAACCTGTCTACAGATGACACACAAAGC	1496
Db	1227	AGGCTAAAAACAGAGGATGAAAGATCCAAATTAACCTGTCTACAGATGACACACAAAGC	1286
Qy	1497	TTCTTCAAGAACATAATTAATGCAATTGAAACAAATTTGAGGAACTAACAGAAAGAGTCAG	1556
Db	1287	TTCTTCAAGAACATAATTAATGCAATTGAAACAAATTTGAGGAACTAACAGAAAGAGTCAG	1346
Qy	1557	AAAAAGTGGACAGGCGAGTCTGAAGGAACTGAGTGAAGAACTGGAACTGGCAGAGAAGG	1616
Db	1347	AAAAAGTGGACAGGCGAGTCTGAAGGAACTGAGTGAAGAACTGGAACTGGCAGAGAAGG	1406
Qy	1617	CTCTGGCTTCCAAAACAGCTGCAAAATGGAATGAAGCAAAACATTTGCCAAGCAGGAAG	1676
Db	1407	CTCTGGCTTCCAAAACAGCTGCAAAATGGAATGAAGCAAAACATTTGCCAAGCAGGAAG	1466
Qy	1677	AGGACCTGGAAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTC	1736
Db	1467	AGGACCTGGAAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTC	1526
Qy	1737	ATGCTGAAGAGCAGCGAGAGAGAAAAATTCATGAGGAAAGGAGCAACTGGCAATTGCAGC	1796
Db	1527	ATGCTGAAGAGCAGCGAGAGAGAAAAATTCATGAGGAAAGGAGCAACTGGCAATTGCAGC	1586
Qy	1797	TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGAGGAGGAGGCTTGTGATGG	1856
Db	1587	TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGAGGAGGAGGCTTGTGATGG	1646
Qy	1857	AGATGACAGACTCGTCATGGGCGAGAACAGTGAATGCTGACCGAGCAGGCTTACCTGTTTC	1916
Db	1647	AGATGACAGACTCGTCATGGGCGAGAACAGTGAATGCTGACCGAGCAGGCTTACCTGTTTC	1706
Qy	1917	AAAGAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTTCCGATTCATTCCTGCC	1976
Db	1707	AAAGAGGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAATATTTCCGATTCATTCCTGCC	1766
Qy	1977	CCAAGTGTGAGAGGTTCTGCTGACATGACAGCTTACAGATTACGTTGATGATGCA	2036
Db	1767	CCAAGTGTGAGAGGTTCTGCTGACATGACAGCTTACAGATTACGTTGATGATGCA	1826
Qy	2037	TCATTTAAGTGTGATCTAT	2056
Db	1827	TCATTTAAGTGTGATGAT	1846

RESULT 4			
LOCUS	CR859573	3265 bp	mRNA
DEFINITION	Pongo pygmaeus mRNA; cdna DKFp459A134 (from clone DKFp459A134).		
ACCESSION	CR859573		
VERSION	CR859573.1		
KEYWORDS	HTC		
SOURCE	Pongo pygmaeus (orangutan)		
ORGANISM	Pongo pygmaeus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.		
REFERENCE	1 (bases 1 to 3265)		
AUTHORS	Ansoorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B., Wewes, H., Well, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.		
CONSRM	The German cDNA Consortium		
TITLE	Submitted (12-NOV-2004)		
JOURNAL	MIPS, Ingolstaedter Landstr.1, D-85764		
COMMENT	Neuberberg, GERMANY		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL European Molecular Biology Laboratories,		

Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp459A134) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.

Please contact RZPD for ordering:

<http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp459A134>

Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

FEATURES

source

```

1..3265
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp459A134"
/tissue_type="cortex"
/clone_lib="459 (synonym: pcor1). Vector pSport1_sfi; host
DH10B; sites SfiI + SfiII"
/dev_stage="adult"
/note="optineurin (Homo sapiens), differentially spliced"
1..3265
/gene="DKFZp459A134"
186..1769
/gene="DKFZp459A134"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH91737.1"
/db_xref="GI:55730024"
/db_xref="UniProtKB/Swiss-Prot:Q5R923"
/translation="MSHQLSLTEKEDSPTEGPPYLALPNLDTFTPEELQQM
KELLTENHOLKMLNNQAMKRFEEISAWTEKQERQFFEIQSKEAKERLMAISH
ENEKLEELKGLKGSERSEDTDDSLRPABAEQEKDOLRTVRLQAEKADLLGI
VSELOLKNSSGSSDSFVEIRMAEGEAGSVKEIKHSPGPTRTVSTSPALSKYRSR
ARGKYNLEHEBLTVSOLLCLREGNQKVERLEIALKKAERVSPFEKASNRSIET
QTEGSTENKDEBEGLTGVSEALNQLVTLFKELOEHAHTLKSEAEMLKKRLQESK
KLTVLQMTTAKEDLEHNTLITLRAQMEVVCSDFAERAAEREKIHKEQLALQALVLL
QNDKQMTTAKEDLEHNTLITLRAQMEVVCSDFAERAAEREKIHKEQLALQALVLL
KENDAFEDGROSLEMEMOSRHGARTSGADQAYLVQRGAEEDRWQRNIRPHSPKC
GEVLDPDIDLIQHVMDCII"
```

ORIGIN

```

Query Match      68.7%; Score 1426.2; DB 6; Length 3265;
Best Local Similarity 90.0%; Pred. No. 0;
Matches 1603; Conservative 0; Mismatches 28; Indels 150; Gaps 1;

QY 297 CAGGAACCTTCGATGTCCATCACTACCTCTCAGCTGCTCACTGAAAAGGAGGACGCC 356
DB 172 CAGGAACCTTCGATGTCCATCACTACCTCTCAGCTGCTCACTGAAAAGGAGGACGCC 231

QY 357 CCAGTGAAAGCACAGGAAATGGACCCGCCACCTGGCCGCCACCAACCTGGACACGTTTA 416
DB 232 CCACGGAAGACAGGAAATGGACCCGCCACCTGGCCGCCACCAACCTGGACACGTTCA 291

QY 417 CCCCGAGGAGCTGTGACGACGATGAAAGAGCTCTGACCAAGAACCAACGAGTGAAG 476
DB 292 CCCCGAGGAGCTGTGACGACGATGAAAGAGCTCTGACCTGAGAACCAACGAGTGAAG 351

QY 477 AAGCCATGAGCTTAATATCAAGCATGAAAGGGAGATTGAGGAGCTTCGGCTGGA 536
DB 352 AAGCCATGAGCTTAATATCAAGCATGAAAGGGAGATTGAGGAGCTTCGGCTGGA 411

QY 537 CAGAGAAACAGAGGAAGAACCGCAGCTTTTGTAGATACAGAGCAAGCAAGCAAGAGAGC 596
DB 412 CAGAGAAACAGAGGAAGAACCGCAGCTTTTGTAGATACAGAGCAAGCAAGAGAGC 471

QY 597 GTCTAATGCGCTTGAGTCAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 656
DB 472 GTCTAATGCGCTTGAGTCAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 531

QY 657 GGAATACAGAAAGGTCATCTGAGGACCCCACTGATGATCTCCAGGCTTCCAGGCGGAG 716
DB 532 GGAATACAGAAAGGTCATCTGAGGACCCCACTGATGATCTCCAGGCTTCCAGGCGGAG 591

QY 717 CGGAGCAGGAAAGGACCAAGCTCAGGACCCAGCGGTGTGTGAGCTTACAGCAGAGAGGCAG 776
```

```

|||||
592 CGGACAGGAAAGGACCCAGCTCAGGACCCAGGTGGTGAAGGCTACAAGCAGAGAGGCAG 651
|||||
777 ACCTGTTGGGCATCGTGTCTGAAGCTCAAGCTGAAGCTCCAGCGGCTCCTCAGAAG 836
|||||
652 ACCTGTTGGGCATCGTGTCTGAAGCTCAAGCTGAAGCTCCAGCGGCTCCTCAGAAG 711
|||||
837 ATTCTCTTTGTTGAAATAGGATGGCTGAAGGAGAAAGCAGAGAGGGTCACTAAAAGAATCA 896
|||||
712 ATTCTCTTTGTTGAAATAGGATGGCTGAAGGAGAAAGCAGAGAGGGTCACTAAAAGAATCA 771
|||||
897 AGCATAGTCTGGGGCCACGAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGA 956
|||||
772 AGCATAGTCTGGGGCCACGAGAACAGTCTCCACTAGCAGGCGCATTTGTCTAAATATAGGA 831
|||||
957 GCAGATCTGCAGATGGGGCCACAGAAATTACTTTCGAACATGAGGAGTTAACTGTGAGCCAGC 1016
|||||
832 GCAGATCTGCAGATGGGGCCACAGAAATTACTTTCGAACATGAGGAGTTAACTGTGAGCCAGC 891
|||||
1017 TCCTCTGTGCTTAAGGGAAGGGAATCAGAAGTGGAGAGACTTGAAGTTGCACCTCAAGG 1076
|||||
892 TCCTCTGTGCTTAAGGGAAGGGAATCAGAAGTGGAGAGACTTGAAGTTGCACCTCAAGG 951
|||||
1077 AGGCCAAAGAAAGAGTTCAGATTTTGAAGAAAGAAACAAGTAATCGTTCTGAGATTGAAA 1136
|||||
952 AGGCCAAAGAAAGAGTTCAGATTTTGAAGAAAGAAACAAGTAATCGTTCTGAGATTGAAA 1011
|||||
1137 CCACAGCAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTG 1196
|||||
1012 CCACAGCAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTG 1071
|||||
1197 GAAGCGAAGTGAAGCAGTGAACCTCCAGGTGCACATCTCTGTTTAAGGAGCTTCAAGAGG 1256
|||||
1072 GAAGCGAAGTGAAGCAGTGAACCTCCAGGTGCACATCTCTGTTTAAGGAGCTTCAAGAGG 1131
|||||
1257 CTCATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAAGAGACTTCAAGAAAAAGTGCAGG 1316
|||||
1132 CTCATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAAGAGACTTCAAGA----- 1180
|||||
1317 CCCTTGAAGGAAATTTCTGCAATTCATCAGATTGATGAAAGCAAGAGCTTGTTC 1376
|||||
1181 ----- 1180
|||||
1377 ATACTACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAAAATCAAAAATGGAAC 1436
|||||
1181 ----- 1180
|||||
1437 AGGCTAAAACAGAGGATGAAAAGTCCAAATTAACTGTGCTACAGATGACACACAAGC 1496
|||||
1181 -----AAAGTCCAAATTAATCTGCTACAGATGACACACAAGC 1221
|||||
1497 TTCTTCAAGAACATTAATGCTATGAAACCAATTGAGCAACTAACCAAGAAAGAGTTCAG 1556
|||||
1222 TTCTTCAAGAACATTAATGCTATGAAACCAATTGAGCAACTAACCAAGAAAGAGTTCAG 1281
|||||
1557 AAAAAGTGGACAGGCGAGTGTGAAAGGAACCTCAGTGAAGAAAACTGGAACCTGGCAGAGAAAG 1616
|||||
1282 AAAAAGTGGACAGGCGAGTGTGAAAGGAACCTGAGTGAAGAAAACTGGAACCTGGCAGAGAAAG 1341
|||||
1617 CTCTGCTTCCAAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAGCAGGAAG 1676
|||||
1342 CTCTGCTTCCAAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAGCAGGAAG 1401
|||||
1677 AGGACTCGAAACCATGACCATCTCAGGCTCAGATGCAAGTTTACTGTTCTGATTTTC 1736
|||||
1402 AGGACTCGAAACCATGACCATCTCAGGCTCAGATGCAAGTTTACTGTTCTGATTTTC 1461
|||||
1737 ATGCTGAAAGAGCAGCAGAGAGAAAAATTCATGAGAAAAAGGAGCAACTGGCATTGCAGC 1796
|||||
1462 ATGCTGAAAGAGCAGCAGAGAGAAAAATTCATGAGCAAAAGGAGCAACTGGCATTGCAGC 1521
|||||
1797 TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTCGAGACCGGAGCAGGCTCCTTGTATGG 1856
|||||
```



```
Db 1201 AATAATGCATTGAAAACAATTGAGGAACCTAACAGAAAAGAGCTCAGAAAAAGTGGACAGG 1260
QY 1571 CGAGTGCTCAAGGAAGTGAAGAACTGGAACCTGGCAGAGAAAGCTCTGGCTTCCAAA 1630
Db 1261 CGAGTGCTGAAGGAAGTGAAGAACTGGAACCTGGCAGAGAAAGCTCTGGCTTCCAAA 1320
QY 1631 CAGCTGCAAAATGCATGAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGGACCTGGAACCC 1690
Db 1321 CAGCTGCAAAATGCATGAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGGACCTGGAACCC 1380
QY 1691 ATGACATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCA 1750
Db 1381 ATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCA 1440
QY 1751 GCGAGAGAGAAATTCATGAGAAAGGACCACTGGCAATTCAGCTGGCAGTTCTGCTG 1810
Db 1441 GCGAGAGAGAAATTCATGAGAAAGGACCACTGGCAATTCAGCTGGCAGTTCTGCTG 1500
QY 1811 AAGAGAAATGATCTTTTGAAGACGAGGAGCGAGTCTTCTGATGGAGATGACAGAGTCT 1870
Db 1501 AAGAGAAATGATCTTTTGAAGACGAGGAGCGAGTCTTCTGATGGAGATGACAGAGTCT 1560
QY 1871 CATGGGCGAGAAACAAAGTGACTCTGACACGAGGCTTACCTTGTTCAAAGAG 1922
Db 1561 CATGGGCGAGAAACAAAGTGACTCTGACACGAGGCTTACCTTGTTCAAAGAG 1612
```

RESULT 6

```

BX384735/c
LOCUS
DEFINITION
  BX384735 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
ACCESSION
  BX384735
VERSION
  BX384735.2 GI:46573611
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Organism
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
  1. (bases 1 to 1089)
  Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  Full-length cDNA libraries and normalization
  Unpublished (2001)
  On May 8, 2003 this sequence version replaced gi:3049275.
COMMENT
  Contact: Genoscope
  Genoscope - Centre National de Sequencage
  2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
  Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
  1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
  end enriched, double-strand cDNA was digested with Not I and cloned
  into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
  was normalized. Library was constructed by Life Technologies, a
  division of Invitrogen. This sequence belongs to sequence cluster
  5584.f
  For more information about this cluster, see
  http://www.genoscope.cns.fr/cdna?c=CS0DL001BE06NP1k=5584.f.
```

FEATURES

```

Location/Qualifiers
  1..1089
    /organism="Homo sapiens"
    /mol_type="RNA"
    /db_xref="taxon:9606"
    /clone="CS0DL001Y112"
    /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
    /cell_line="RAMOS CELL LINE"
    /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
    25-NORMALIZED"
    /note="1st strand cDNA was primed with a NotI-oligo(dT)
    primer. Five prime end enriched, double-strand cDNA was
    digested with Not I and cloned into the Not I and EcoR V
    sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```

Query Match 47.5%; Score 986.6; DB 4; Length 1089;
Best Local Similarity 96.1%; Pred. No. 5,1e-244;
Matches 1016; Conservative 14; Mismatches 24; Indels 3; Gaps 2;
QY 1014 AGCTCTCTGTGCTTAAGGGAAGGAATCAGAAAGGTGGAGAGACTTGAAGTTGCACTCA 1073
Db 1054 ACCCGCYCTSTKTCCTTAGGGARGGAATCAGAAAGTG- -ARAAAATKGAAGTGCACTCA 997
QY 1074 AGGAGCCCAAGAAAGAGTTTCAGATTTTGAAGAGAAAACACAGTAATCGTCTTCGATGTTG 1133
Db 996 ARGAGCCCAAGAAAGAGTTTCAGATTTTGAAGAGAAAACAAAGTAATCGTCTTCGATGTTG 937
QY 1134 AAACCCACAGACAGAGGGGAGCAGAGAGAAAGAGAAATGATGAAGAGAAAAGCCCGGAGACTG 1193
Db 936 AAWCCACAGACAGAGGGGAGCAGAGAGAAAGAGAAATGATGAAGAGAAAAGCCCGGAGACTG 877
QY 1194 TTGGAAGCGAAGTGAAGACACTGAACCTCCAGGTGCACATCTCTGTTTAAGGAGCTTCAAG 1253
Db 876 TTGGAAGCGAAGTGAAGCACTGAACCTCCAGGTGCACATCTCTGTTTAAGGAGCTTCAAG 817
QY 1254 AGGCTCATCAAAACTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAGTGC 1313
Db 816 AGGCTCATCAAAAACTCGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAGTGC 757
QY 1314 AGGCCCTTTGAAGGAAAAAATCTGCAATTCATCAGAGTTGAATGAAAAGCAAGAGCTTGG 1373
Db 756 AGGCCCTTTGAAGGAAAAAATCTGCAATTCATCAGAGTTGAATGAAAAGCAAGAGCTTGG 697
QY 1374 TTTATATCAAAAGTTAGAGCTTACAAGTGGAAAAGCATGCTATCAGAAAATCAAAATGG 1433
Db 696 TTTATATCAAAAGTTAGAGCTTACAAGTGGAAAAGCATGCTATCAGAAAATCAAAATGG 637
QY 1434 AACAGGCTTAAACACAGAGGATGAAAAGTCCAAATTAACCTGCTCAGAGTACACACAACA 1493
Db 636 HACAGGCTTAAACACAGAGGATGAAAAGTCCAAATTAACCTGCTCAGAGTAAACACAACA 577
QY 1494 AGCTTCTTCAAGAACATAATTAATGATGAAAACAAATTCAGGAACTAAACAAGAAAGAGT 1553
Db 576 AGCTTCTTCAAGAACATAATTAATGATGAAAACAAATTCAGGAACTAAACAAGAAAGAGT 517
QY 1554 CAGAAAAAGTGAAGAGGCGAGTGTGAAGGAACCTGAGTGAAGAAAACCTGGAACCTGGCAGAGA 1613
Db 516 CAGAAAAAGTGGACAGAGGCGAGTGTGAAGGAACCTGAGTGAAGAAAACCTGGAACCTGGCAGAGA 457
QY 1614 AGGCTCTGCTTCCAAAACAGCTGCAAAATGGAATGAAGCAAAACCATTTGCCAAGCAGG 1673
Db 456 AGGCTCTGCTTCCAAAACAGCTGCAAAATGGAATGAAGCAAAACCATTTGCCAAGCAGG 397
QY 1674 AAGAGGACCTGGAAACCACTGACCATCTCCAGGGCTCAGATGGAAGTTTACTGTTCTGATT 1733
Db 396 AAGAGGACCTGGAAACCACTGACCATCTCCAGGGCTCAGATGGAAGTTTACTGTTCTGATT 337
QY 1734 TTCAATGCTGAAGACAGCGAGAGAGAAAATTCATGAGAAAAGAGCAACTGGCAATTGC 1793
Db 336 TTCAATGCTGAAGACAGCGAGAGAGAAAATTCATGAGAAAAGAGCAACTGGCAATTGC 277
QY 1794 AGCTGGCAGTTCTGCTGAAAGAGAAATGATGCTTTCCGAAGACGAGCGAGCTCCTTGA 1853
Db 276 AGCTGGCAGTTCTGCTGAAAGAGAAATGATGCTTTCCGAAGACGAGCGAGCTCCTTGA 217
QY 1854 TGGAGATGACAGTCTCATGTTGGGGCGAGAAACAAAGTGATCTCTGACACGAGGCTTACCTTGG 1913
Db 216 TGGAGATGACAGTCTCATGTTGGGGCGAGAAACAAAGTGATCTCTGACACGAGGCTTACCTTGG 157
QY 1914 TTCAAGAGGAGCTGAGGACAGGGA CTGGCGGCAACAGCGGAATATTTCGATTTCATTCCT 1973
Db 156 TTCAAGAGGAGCTGAGGACAGGTA CTGGCGGCAACAGCGGAATATTTCGATTTCATTCCT 97
QY 1974 GCCCAAGTGTGGAGAGGTTCTGCTCTGACATAGACAGTTTACAGATTTCACGTGATGGATT 2033
Db 96 GCCCAAGTGTGGAGAGGTTCTGCTCTGACATAGACAGTTTACAGATTTCACGTGATGGATT 37
```

```
QY 2034 GCATCATTTAAGTGTGATATACACCTCCCAAAAC 2070
|||||
Db 36 GCATCATTTAAGTGTGATGT-TCACCTCCCAAAAC 1

RESULT 7
AL583163/c
LOCUS
DEFINITION
AL583163 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL009YF13 3-PRIME, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 1068)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12951861.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremonieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5584.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?sc=CS0DL009CC07NP1&c=5584.f.
FEATURES
source
Location/Qualifiers
1..1068
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YF13"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."

Query Match 47.3%; Score 982.6; DB 1; Length 1068;
Best Local Similarity 97.9%; Pred. No. 5.6e-243;
Matches 1019; Conservative 7; Mismatches 11; Indels 4; Gaps 3;

QY 1002 TAACTGTGACCCAGCTCTCTGTGCTTAAGGAGGGGAATCAGAGGTGG-AGAGACTT 1060
|||
Db 1041 TTAAYKTGACCGCGCTCTGTGCTTAAGGAGGGGAATCAGAGGTGGAGAGACTT 982
|||
QY 1061 GAAGTTGC--ACTCAGGAGGCCAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTA 1118
|||
Db 981 GAAAGTGCCACTCCAAGGAGSCCAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTA 922
|||
QY 1119 ATCGTTCTGAGTTGAACCCACAGAGGGGAGCAGAGGAAGAAATGATGAAGAGA 1178
|||
Db 921 ATCGTTCTGAGTTGAACCCACAGAGGGGAGCAGAGGAAGAAATGATGAAGAGA 862
|||
QY 1179 AAGCCCGGAGACTGTTGAAGCCCAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGT 1238
|||
Db 861 AAGCCCGGAGACTGTTGAAGCCCAAGTGAAGCACTGAACCTCCAGGTGAAAWCTCTGT 802
|||
QY 1239 TTAAGGAGCTTCAAGAGGCTCATAC-AAACTCAGCGAAGCTCAGCTTAATGAAGAGAGA 1297
|||||
```

```
Db 801 TTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAGAGA 742
QY 1298 CTTCAAGAAAGTGTCTAGGCTCTTGAAGGAAATTTCTCAATTTCCATCAGAGTTCAAT 1357
|||||
Db 741 CTTCAAGAAAGTGTCTAGGCTCTTGAAGGAAATTTCTCAATTTCCATCAGAGTTCAAT 682
|||||
QY 1358 GAAAGCAAGAGCTTGTATTACTAAACAAAAGTTAGAGCTACAAGTGGAAAAGCATGCTA 1417
|||||
Db 681 GAAAGCAAGAGCTTGTATTACTAAACAAAAGTTAGAGCTACAAGTGGAAAAGCATGCTA 622
|||||
QY 1418 TCAGAAATCAAAATGGAACAGGCTTAAACACAGAGGATGAAAAGTCCAAATTAACCTGCTA 1477
|||||
Db 621 TCAGAAATCAAAATGGAACAGGCTTAAACACAGAGGATGAAAAGTCCAAATTAACCTGCTA 562
|||||
QY 1478 CAGATGACACACAAACAGCTTCTTCAAGAACATAAATCAATTTGAAAAACAATTTGAGGAA 1537
|||||
Db 561 CAGATGACACACAAACAGCTTCTTCAAGAACATAAATCAATTTGAAAAACAATTTGAGGAA 502
|||||
QY 1538 CTAACAAAGAAAGAGTCAGAAAAGTGGACAGGCGAGTGTCTGAAGGAACTGAGTGAAGAAA 1597
|||||
Db 501 CTAACAAAGAAAGAGTCAGAAAAGTGGACAGGCGAGTGTCTGAAGGAACTGAGTGAAGAAA 442
|||||
QY 1598 CTGGAACTGGCAGAGAGGCTCTGGCTTCCAAACAGCTGCAAAATGGATGAATGAAGCAA 1657
|||||
Db 441 CTGGAACTGGCAGAGAGGCTCTGGCTTCCAAACAGCTGCAAAATGGATGAATGAAGCAA 382
|||||
QY 1658 ACCATTGCCAAGCAGGAAGGACCTGGAAACCATGACCATCTCAGGGCTCAGATGCAA 1717
|||||
Db 381 ACCATTGCCAAGCAGGAAGGACCTGGAAACCATGACCATCTCAGGGCTCAGATGCAA 322
|||||
QY 1718 GTTTACTGTTCTGATTTTCAATGCTGAAAGAGCAGCAGAGAGAGAAAATTCATGAGGAAAAG 1777
|||||
Db 321 GTTTACTGTTCTGATTTTCAATGCTGAAAGAGCAGCAGAGAGAGAAAATTCATGAGGAAAAG 262
|||||
QY 1778 GAGCAACTGGCATTGCGAGCTGGCAGTTCTGCTGAAAGAGAGATGATGCTTTGGAAGACGA 1837
|||||
Db 261 GAGCAACTGGCATTGCGAGCTGGCAGTTCTGCTGAAAGAGAGATGATGCTTTGGAAGACGA 202
|||||
QY 1838 GGCAGGAGCTCTTGTGATGAGATGCGAGTCTGTCATGGGGCGAGAACAACTGACTCTGAC 1897
|||||
Db 201 GGCAGGAGCTCTTGTGATGAGATGCGAGTCTGTCATGGGGCGAGAACAACTGACTCTGAC 142
|||||
QY 1898 CAGCAGGCTTACCTTTCAAAGAGAGCTGAGGACAGGGAATCTGGCGGCAACAGCGGAAT 1957
|||||
Db 141 CAGCAGGCTTACCTTTCAAAGAGAGCTGAGGACAGGGAATCTGGCGGCAACAGCGGAAT 82
|||||
QY 1958 ATTCCGATTCATTCCTGCCCCCAAGTGTGAGAGAGTTCTGCTGACATAGACAGTTACAG 2017
|||||
Db 81 ATTCCGATTCATTCCTGCCCCCAAGTGTGAGAGAGTTCTGCTGACATAGACAGTTACAG 22
|||||
QY 2018 ATTCAGCTGATGATGATGATC 2038
|||||
Db 21 ATTCAGCTGATGATGATGATC 1

AL570028 1032 bp mRNA linear EST 05-APR-2004
AL570028 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1003YG24 3-PRIME, mRNA sequence.
ACCESSION AL570028
VERSION AL570028.3 GI:46236271
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 1032)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:31291459.
```


Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5584.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI003BD12NP1&c=5584.f.

FEATURES

Location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI003YG24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 46.1%; Score 957.6; DB 1; Length 1032;
Best Local Similarity 95.6%; Pred. No. 1.7e-236; Mismatches 14; Indels 1; Gaps 1;
Matches 978; Conservative 14; Mismatches 30; Indels 1; Gaps 1;

QY	1035	AAGGGAATCAGAGGTGGAGAGACTTGAAGTTCACCTCAAGGAGGCCAAGAAAGAGTTT	1094
DB	1023	AGGRAGGATCAGAGGGAGAGACTTGAATGCACTCARGGGCCAAAGAAAGAGTTM	964
QY	1095	CAGATTTTGAAGAAGAAACAAGTAATCGTTCTGAGATTGAAACCCACAGAGGGGAGCA	1154
DB	963	CAGATTTTGAAGAAGAAACAAGTAATCGTTCTGAGATTGAAACCCACAGAGGGGAGCA	904
QY	1155	CAGAGAAGAGATGATGAAGAGAAAGGCCCGGAGACTCTGGAAGCGAAGTGGAGACAC	1214
DB	903	CAGAGAAGAGATGATGAAGAGAAAGGCCCGGAGACTCTGGAAGCGAAGTGGAGACAC	844
QY	1215	TGAACCTCCAGGTGACATCTCTGTTTAAAGAGGCTTCAAGAGGCTCATACAAAACCTCAGCG	1274
DB	843	TGAACCTCCAGGTGACATCTCTGTTTAAAGAGGCTTCAAGAGGCTCATACAAAACCTCAGCG	784
QY	1275	AAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAGTGTGAGGCTTTGAAAGGAAATTT	1334
DB	783	AAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAGTGTGAGGCTTTGAAAGGAAATTT	724
QY	1335	CTGCAATTCATCAGAGTTGAATGAAGCAAGAGCTTCTTATATCTAACAAGAAAGTTAG	1394
DB	723	CTGCAATTCATCAGAGTTGAATGAAGCAAGAGCTTCTTATATCTAACAAGAAAGTTAG	664
QY	1395	AGCTACAGTGAAGAGCTGCTATCAGAAATCAAAATGAAACAGGCTAAACAGAGGATG	1454
DB	663	AGCTACAGTGAAGAGCTGCTATCAGAAATCAAAATGAAACAGGCTAAACAGAGGATG	604
QY	1455	AAAAGTCCAAATTAATCTGTGCTACAGATGACACAAAGCTTCTTCAAGAACATAATA	1514
DB	603	AAAAGTCCAAATTAATCTGTGCTACAGATGACACAAAGCTTCTTCAAGAACATAATA	544
QY	1515	ATGCATTGAAACAATTTGAGGAACTAACAGAAAGAGTTCAGAAAGTGGACAGGGCAG	1574
DB	543	ATGCATTGAAACAATTTGAGGAACTAACAGAAAGAGTTCAGAAAGTGGACAGGGCAG	484
QY	1575	TGCTGAAGGAACCTGAGTGAAGAACTGGAACCTGSCAGAGAGGCTCTGGCTTCCAAACAGC	1634
DB	483	TGCTGAAGGAACCTGAGTGAAGAACTGGAACCTGSCAGAGAGGCTCTGGCTTCCAAACAGC	424
QY	1635	TGCAATGATGAATGAAGCAACCACTTGCACAGAGAGAGGAGCTTGGAAACCATCA	1694
DB	423	TGCAATGATGAATGAAGCAACCACTTGCACAGAGAGAGGAGCTTGGAAACCATCA	364

QY	1695	CCATCTCAGGGCTCAGATGGAAGTTTACTGTCTCTGATTTTTCATGCTGAAAGAGACGCGA	1754
DB	363	CCATCTCAGGGSTCAGATGGAAGTTTACTGTCTCTGATTTTTCATGCTGAAAGAGACGCGA	304
QY	1755	GAGAGAAATTCATGAGGAAAAGGAGCAATGCGCATTCAGCTGCGAGTTCTGCTGAAAG	1814
DB	303	GAGAGAAATTCATGAGGAAAAGGAGCAATGCGCATTCAGCTGCGAGTTCTGCTGAAAG	244
QY	1815	AGAAATGATGCTTTCGAAGACGGAGGAGGAGCTCTTGAATGAGATGAGAGTCGTCATG	1874
DB	243	AGAAATGATGCTTTCGAAGACGGAGGAGGAGCTCTTGAATGAGATGAGAGTCGTCATG	184
QY	1875	GGGCGAGAAACAGTCACTCTGACACGAGGAGCTTACCTGTTCAAAGAGGAGCTGAGGACA	1934
DB	183	GGGCGAGAAACAGTCACTCTGACACGAGGAGCTTACCTGTTCAAAGAGGAGCTGAGGACA	124
QY	1935	GGGACTGGCGCAACAGCGGAATATTCGATTCATTCCTGCCCCCAAGTGTGAGAGGTTTC	1994
DB	123	GGGACTGGCGCAACAGCGGAATATTCGATTCATTCCTGCCCCCAAGTGTGAGAGGTTTA	64
QY	1995	TGCCTGCATACAGACAGTTACAGATTCACGTCATGATGATGATGATGATGATGATGATG	2053
DB	63	TGCCTGCATACAGACAGTTACAGATTCACGTCATGATGATGATGATGATGATGATGATG	4
QY	2054	TAT 2056	
DB	3	TAT 1	

RESULT 9
AL562664/c
LOCUS
DEFINITION
AL562664 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC016YG04 3-PRIME, mRNA sequence.
ACCESSION
AL562664
VERSION
AL562664.3 GI:46227675
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1. (bases 1 to 1021)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On Feb 15, 2001 this sequence version replaced gi:31286674.
COMMENT
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5584.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DC016BD02NP1&c=5584.f.

FEATURES
Location/Qualifiers
1..1021
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC016YG04"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match	45.1%;	Score 937.2;	DB 1;	Length 1021;	
Best Local Similarity	97.4%;	Pred. No. 3.2e-231;			
Matches 963;	Conservative 5;	Mismatches 19;	Indels 2;	Gaps 2;	
Qy	1074	AGGAGGCCAAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAAAGTAATCGTCTCAGAGATTG	1133		
Db	992	AAGRGCCAAAGAAAGHWTTCAG-WTTTGAAGAAAGAAACAAAGTAATCGTCTCAGAGATTG	934		
Qy	1134	AAACCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTG	1193		
Db	933	AAACCCARACAGA-GGGAGCACAGAGAAAGAGATGATGAAGAGAAAGGCCCGGAAACTG	875		
Qy	1194	TTGNAAGCGAGTGGAGCACTGAACCTCCAGGTGACATCTCTTTTAAAGAGCTTCAAG	1253		
Db	874	TTGNAAGCGAAGTGGAGCACTGAACCTCCAGGTGACATCTCTTTTAAAGAGCTTCAAG	815		
Qy	1254	AGGCTCATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAGAGAGACTTCAAGAAAAGTGC	1313		
Db	814	AGGCTCATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAGTGC	755		
Qy	1314	AGGCCCTTGAAGGAAATAATCTGCAATTTCAATCAGAGTTGAATGAAGAAAGCAAGAGCTTG	1373		
Db	754	AGGCCCTTGAAGGAAATAATCTGCAATTTCAATCAGAGTTGAATGAAGAAAGCAAGAGCTTG	695		
Qy	1374	TTTATACTACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGG	1433		
Db	694	TTTATACTACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGG	635		
Qy	1434	AACAGGCTAAAACAGAGAGATGAAAGTCCAAATTTAACTGTCTACAGATGACACACAACA	1493		
Db	634	AACAGGCTAAAACAGAGAGATGAAAGTCCAAATTTAACTGTCTACAGATGACACACAACA	575		
Qy	1494	AGCTTCTTCAAGAACATAATAATGCATTTGAAACAAATTTGAGAACTTAACAGAAAAGAGT	1553		
Db	574	AGCTTCTTCAAGAACATAATAATGCATTTGAAACAAATTTGAGAACTTAACAGAAAAGAGT	515		
Qy	1554	CAGAAAAGTGCACAGGGCAGTCTGAAGGAAGTCTGAGTGAAGAACTGGAAGTGCAGAGA	1613		
Db	514	CAGAAAAGTGCACAGGGCAGTCTGAAGGAAGTCTGAGTGAAGAACTGGAAGTGCAGAGA	455		
Qy	1614	AGGCTCTGGCTTCCAAACAGCTGCAAAATGGAATGAATGAAGCAACCATTTGCCAAGCAGG	1673		
Db	454	AGGCTCTGGCTTCCAAACAGCTGCAAAATGGAATGAATGAAGCAACCATTTGCCAAGCAGG	395		
Qy	1674	AAGAGACCTTGAACACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATT	1733		
Db	394	AAGAGACCTTGAACACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATT	335		
Qy	1734	TTCAATGCTGAAAGAGCAGAGAGAGAAATTTCAATGAGGAAAGAGCAACTGCGATTGC	1793		
Db	334	TTCAATGCTGAAAGAGCAGAGAGAGAAATTTCAATGAGGAAAGAGCAACTGCGATTGC	275		
Qy	1794	AGCTGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAGGAGCAGTCTCTTGA	1853		
Db	274	AGCTGGCAGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGACGGAGGAGCAGTCTCTTGA	215		
Qy	1854	TGGAGATGACAGAGTCGTATGGGGCGAGAACAGTGAATCTGACAGCAGGCTTACCTTG	1913		
Db	214	TGGAGATGACAGAGTCGTATGGGGCGAGAACAGTGAATCTGACAGCAGGCTTACCTTG	155		
Qy	1914	TTCAAAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAAATATTTCCGATTCATTCCCT	1973		
Db	154	TTCAAAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAAATATTTCCGATTCATTCCCT	95		
Qy	1974	GCCCCAAGTGTGGAGAGGTTCTGCTGCACATAGACACGTTTACAGATTCACGTGATGAT	2033		
Db	94	GCCCCAAGTGTGGAGAGGTTCTGCTGCACATAGACACGTTTACAGATTCACGTGATGAT	35		
Qy	2034	GCATCATTTAAGTGTGATGTATCACCTC	2062		
Db	34	GCATCATTTAAGTGTGATGTATTCCTCCTCCC	6		

Db	368	GTCTAATGGCCTTGAGTCATGAGAATGAGAAATGAGGAAGAGCTTGAAAACTAAAAAG	427
QY	657	GGAATATCAGAAAGTCATCTGAGGACCCACCTGATGACTCCAGGCTTCCAGGCGCGAG	716
Db	428	GGAATATCAGAAAGTCATCTGAGGACCCACCTGATGACTCCAGGCTTCCAGGCGCGAG	487
QY	717	CGGAGCAGGAAAAAGGACCTCAGGACCCAGCTGAGGCTCAAGCAGAGAAAGGCGAG	776
Db	488	CGGAGCAGGAAAAAGGACCTCAGGACCCAGCTGAGGCTCAAGCAGAGAAAGGCGAG	547
QY	777	ACCTGTGGGCATCTGTGTCTGAATCTGACCTCAAGCTCAAGCTCCAGCGGCTCCTCAGAAG	836
Db	548	ACCTGTGGGCATCTGTGTCTGAATCTGACCTCAAGCTCAAGCTCCAGCGGCTCCTCAGAAG	607
QY	837	ATTCCTCTTGTCTGAAATTAGGATGGCTGAGGAGAGAGCAAGGCTCAGTAAAGAAATCA	896
Db	608	ATTCCTCTTGTCTGAAATTAGGATGGCTGAGGAGAGAGCAAGGCTCAGTAAAGAAATCA	667
QY	897	AGCATAGTCCTGGGCCACGAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGA	956
Db	668	AGCATAGTCCTGGGCCACGAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGA	727
QY	957	CGAGATCTCGAGATGGGCGCAAGAAATTAATTCGAAATGAGGAGTAACTGTGAGCCAGC	1016
Db	728	CGAGATCTCGAGATGGGCGCAAGAAATTAATTCGAAATGAGGAGTAACTGTGAGCCAGC	787
QY	1017	TCTCTGTGCTTAAGGGAAGGGAATCAGAAGGTGGAGAGCTTGAAGTTCACCTCAAGG	1076
Db	788	TCTCTGTGCTTAAGGGAAGGGAATCAGAAGGTGGAGAGCTTGAAGTTCACCTCAAGG	847
QY	1077	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAACTCTTCAGATTTGAAA	1136
Db	848	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAACTCTTCAGATTTGAAA	907
QY	1137	CCCAGACAGGGGACACAGAGAAAGAAATGATGAAGAGAAAGCCCGGAGAGCTGTG	1196
Db	908	CCCAGACAGGGGACACAGAGAAAGAAATGATGAAGAGAAAGCCCGGAGAGCTGTG	967
QY	1197	GAAGCGA 1203	
Db	968	GAAGCGA 974	

RESULT 11
BX343673/c
LOCUS
DEFINITION
BX343673 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1009YB18 3-PRIME, mRNA sequence.
ACCESSION
BX343673
VERSION
BX343673.2 GI:46283489
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1012)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30347281.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5584.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS1A1003B10NP1&c=5584.f.

FEATURES	Location/Qualifiers	
Source	1. 1012	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="CS0D1009YB18"	
	/issue_type="PLACENTA COT 25-NORMALIZED"	
	/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"	
	/note="1st strand cDNA was primed with a NotI-oligo(dT)	
	primer. Five prime end enriched, double-strand cDNA was	
	digested with Not I and cloned into the Not I and EcoR V	
	sites of the pCMVSPORT 6 vector. Library was normalized."	
ORIGIN		
	Query Match 42.5%; Score 882.4; DB 4; Length 1012;	
	Best Local Similarity 92.4%; Pred. No. 5.2e-217;	
	Matches 905; Conservative 27; Mismatches 45; Indels 2; Gaps 2;	
QY	1086 AAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCTGTTCTGAGATTTGAAACCCAGACAG	1145
Db	977 AAAGAGGKTTTCMGATTTTGAAGAAACAAADTATBTCTGTGATK-AAACCCAGACAG	919
QY	1146 AGGGAGCACAGAGAAAGAGAAATGATGAAGAAAGGCGCGAGACTGTGGAAGCGAAG	1205
Db	918 A-GGGRGCACAGAGAAAGAGAAATGATGAAGAAAGGSSSGAGACTGTGGAAGCGAAG	860
QY	1206 TGGAGGACTCAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAA	1265
Db	859 TGGAGGACTCAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAA	800
QY	1266 AACTCAGCGAAGCTGAGCTAATGAAGAAAGAGACTTCAAGAAAGTGTTCAGGGCCTTGA	1325
Db	799 AAACCTCRSGAAGCTGAGCTAATGAAGAAAGAGACTTCAAGAAAGTGTTCAGGGCCTTGA	740
QY	1326 GGAAAAATTCGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTGTTATATACAA	1385
Db	739 GGAAAAATTCGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTGTTATATACAA	680
QY	1386 AAAAGCTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAGAGGCTTAA	1445
Db	679 AAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAGAGGCTTAA	620
QY	1446 CAGAGGATGAAAAAGTCCAAATTTAACTGTGCTACAGATGACACACAAAGCTTCTTCAAG	1505
Db	619 CAGAGGATGAAAAAGTCCAAATTTAACTGTGCTACAGATGACACACAAAGCTTCTTCAAG	560
QY	1506 AACATAATATGCAATTTGAAAAACAATGAGGAACCTAACAGAAAGAGTCAAGAAAGTGG	1565
Db	559 AACATAATATGCAATTTGAAAAACAATGAGGAACCTAACAGAAAGAGTCAAGAAAGTGG	500
QY	1566 ACAGGGCAGTCTGAAAGGAACCTGAGTGAAGAAACTCGAACTGGCAGAGAGGCTCTGGCTT	1625
Db	499 ACAGGGCAGTCTGAAAGGAACCTGAGTGAAGAAACTCGAACTGGCAGAGAGGCTCTGGCTT	440
QY	1626 CCAAAACAGCTGCAAAATGGATGAAATGAAAGCAAAACCTTCCCAAGCAGGAAGAGGACCTGG	1685
Db	439 CCAAAACAGCTGCAAAATGGATGAAATGAAAGCAAAACCTTCCCAAGCAGGAAGAGGACCTGG	380
QY	1686 AAACCATGACCATCTCCAGGCTCAGATGGAAGATTTACTGTTCTGATTTTTCATGCTGAAA	1745
Db	379 AAACCATGACCATCTCCAGGCTCAGATGGAAGATTTACTGTTCTGATTTTTCATGCTGAAA	320
QY	1746 GAGCAGCGAGAGAGAAAAATTCATGAGGAAAAAGGAGCAACTGGCATTGCAAGCTGGCAGTTTC	1805
Db	319 GAGCAGCGAGAGAGAAAAATTCATGAGGAAAAAGGAGCAACTGGCATTGCAAGCTGGCAGTTTC	260
QY	1806 TGCTGAAAGAGAAATGATGCTTTGGAAGCGGAGGAGGAGGCTTCTGATGGAAGATCAGA	1865
Db	259 TGCTGAAAGAGAAATGATGCTTTGGAAGCGGAGGAGGAGGCTTCTGATGGAAGATCAGA	200
QY	1866 GTCTGATGGGCGAGAACAGTGTCTGACACAGGCTTACCTGTTTCAAGAGGAG	1925
Db	199 GTCTGATGGGCGAGAACAGTGTCTGACACAGGCTTACCTGTTTCAAGAGGAG	140


```
Db 119 CAGGAACCTTCGCAATGTCCCACTCAACCTCTCAGCTGCCTCACTGAAAAGGAGCAGC 178
Qy 357 CCAGTGAAGACACAGGAATGGACCCGCCACCTGGCCACCCCAAACTGGACACGTTTA 416
Db 179 CCAGTGAAGACACAGGAATGGACCCGCCACCTGGCCACCCCAAACTGGACACGTTTA 238
Qy 417 CCCCGAGGAGCTCTGTCAGCAGATGAAGAGCTCTTGACCAAGAACACACAGCTGAAG 476
Db 239 CCCCGAGGAGCTCTGTCAGCAGATGAAGAGCTCTTGACCAAGAACACACAGCTGAAG 298
Qy 477 AAGCCATGAAGCTAAATCAAGCCATGAAGGAGATTTGAGGAGCTTTCCGCTTGA 536
Db 299 AAGCCATGAAGCTAAATCAAGCCATGAAGGAGATTTGAGGAGCTTTCCGCTTGA 358
Qy 537 CAGAGAAACAGAGGAAGACGCCAGTGTGATGATACAGAGCAAGAAAGCAAGAGC 596
Db 359 CAGAGAAACAGAGGAAGACGCCAGTGTGATGATACAGAGCAAGAAAGCAAGAGC 418
Qy 597 GTCTAATGGCTTGAGTCATGAGAAATGAGAAATGAAGGAGAGCTTTGGAATACTAAAG 656
Db 419 GTCTAATGGCTTGAGTCATGAGAAATGAGAAATGAAGGAGAGCTTTGGAATACTAAAG 478
Qy 657 GGAATCAGAAAGCTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAG 716
Db 479 GGAATCAGAAAGCTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAG 538
Qy 717 CGAGCAGGAAAGGACGAGCTCAGGACCCAGGTTGGTGAGGCTCAAGCAGAGAGGCGAG 776
Db 539 CGAGCAGGAAAGGACGAGCTCAGGACCCAGGTTGGTGAGGCTCAAGCAGAGAGGCGAG 598
Qy 777 ACCTGTTGGGATCGTGTCTGAACTGAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAG 836
Db 599 ACCTGTTGGGATCGTGTCTGAACTGAGCTCAAGCTCAAGCCGAACTCCAGCGGCTCTCTCAGAAG 658
Qy 837 ATTCCTTTGTTGAATTAGATGCTGAGGAGGAGCAGAGAGGTCAGTAAAGAAATCA 896
Db 659 ATTCCTTTGTTGAATTAGATGCTGAGGAGGAGCAGAGAGGTCAGTAAAGAAATCA 718
Qy 897 AGCATAGCTCTGGGCCACAGAAACAGTCTCCAGTGGCAGCGCATTCCTAAATATAGGA 956
Db 719 AGCATAGCTCTGGGCCACAGAAACAGTCTCCAGTGGCAGCGCATTCCTAAATATAGGA 778
Qy 957 GCAGATCTGAGATGGGCCCAAGAAATTAATTCGAAACATGAGGAGTTAACTGTGAGCCAGC 1016
Db 779 GCAGATCTGAGATGGGCCCAAGAAATTAATTCGAAACATGAGGAGTTAACTGTGAGCCAGC 838
Qy 1017 TCCTGCTGCTTAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG 1076
Db 839 TCCTGCTGCTTAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG 898
Qy 1077 AGGCCAAAGAAAGATTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAA 1136
Db 899 AGGCCAAAGAAAGATTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAA 956
Qy 1137 CCCAGACAGAGGGGAGCAGCAGAGAAAGAGATGATGAAGAGAAAGGCC 1184
Db 957 -CCAGACAGAGGGGAGCAC--AAGAAAGAGATGATGAAGAAAGGCC 1001
```

```
RESULT 15
DQ036022
LOCUS
DEFINITION Pan troglodytes OPTN gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION DQ036022
VERSION DQ036022.1
KEYWORDS GI:66887231
SOURCE GSS.
ORGANISM Pan troglodytes (chimpanzee)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
```

```
REFERENCE 1 (bases 1 to 1497)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civallo,D.,
White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL PLOS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1497)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civallo,D.,
White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source 1.1497
gene <1..51497
ORIGIN /locus_tag="HC5124"
Query Match 39.7%; Score 824.4; DB 14; Length 1497;
Best Local Similarity 55.3%; Pred. No. 6.4e-202;
Matches 828; Conservative 0; Mismatches 669; Indels 0; Gaps 0;
Qy 341 GAAAAGGAGGACAGCCCGAGTGAAGACACAGGAAATGGACCCCGACCTGGCCACCCCA 400
Db 1 GAAAAGGAGGACAGCCCGAGTGAAGACACAGGAAATGGACCCCGACCTGGCCACCCCA 60
Qy 401 AACCTGACACGTTTACCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTTGACCAAG 460
Db 61 AACCTGACACGTTTACCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Qy 461 AACCCAGCAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 520
Db 121 AACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
Qy 521 GAGCTTTCGCGCTGGACAGAGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
Qy 581 AAAGAAAGCAAAAGAGCGTCTAATGGCTTTGAGTCTATGAGAAATGAAGAGGAGAG 640
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
Qy 641 CTTGGAAACTAAAGGGGAAATCAGAAAGTCTCTGAGGAGCCCACTGATGATCCAGG 700
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
Qy 701 CTTCCCGAGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 760
Db 361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420
Qy 761 CAAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 820
Db 421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy 821 AGCGGCTCTCAGAGAGTCTCTTTGTTGAAATTTAGGATGCTGAAAGAGGAGGAGGAGG 880
Db 481 AGCGGCTCTCAGAGAGTCTCTTTGTTGAAATTTAGGATGCTGAAAGAGGAGGAGGAGG 540
Qy 881 TCAGTAAAGAAATCAAGCATAGTCTGGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 940
Db 541 TCAGTAAAGAAATCAAGCATAGTCTGGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Qy 941 TTGCTCTAAATATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1000
```


Qy 446 GAGCTCTGACCAAGAACCCAGCTGAAAGAGCCATGAGCTAAATAATCAAGCCATG 505
Db 301 GAGCTCTGACCGAGAACCCAGCTGAAAGAGCCATGAGCTAAATAATCAAGCCATG 360
Qy 506 AAACGGAGATTGAGGAGCTTCCGCTCGGACAGAGAAACAGAGAGGAGAACGCCAGTTT 565
Db 361 AAAGGAGATTGAGGAGCTTCCGCTCGGACAGAGAAACAGAGAGGAGAACGCCAGTTT 420
Qy 566 TTTGAGATACAGAGCAAGAGCAAAAGAGCGCTTAATGGCTTTGAGTCATGAGAATGAG 625
Db 421 TTTGAGATACAGAGCAAGAGCAAAAGAGCGCTTAATGGCTTTGAGTCATGAGAATGAG 480
Qy 626 AAATTGAGAGAGCTTGGAAACTAAAGGAGAAATCAGAAAGCTCATCTGAGGACCC 685
Db 481 AAATTGAGAGAGCTTGGAAACTAAAGGAGAAATCAGAAAGCTCATCTGAGGACCC 540
Qy 686 ACTGATGACTCCAGGCTTCCAGGGCCGAGAGCGAGCAGGAGAAAGGACCGCTCAGGACC 745
Db 541 ACTGATGACTCCAGGCTTCCAGGGCCGAGAGCGAGCAGGAGAAAGGACCGCTCAGGACC 600
Qy 746 CAGGTGCTGAGGCTTACAGAGAGAGGAGAGCAGCTGTTGGGCATCGTGTCTGAACTGCAG 805
Db 601 CAGGTGCTGAGGCTTACAGAGAGAGGAGAGCAGCAGCTGTTGGGCATCGTGTCTGAACTGCAG 660
Qy 806 CTCAGCTGAATCTCAGCGGCTCCTCAGAGATTCCTTTGTTGAAATGAGATGAGTGCAG 865
Db 661 CTCAGCTGAATCTCAGCGGCTCCTCAGAGATTCCTTTGTTGAAATGAGATGAGTGCAG 720
Qy 866 GGAGAGCAGAGGCTCAGTAAGAGAAATCAAGCATAGTCTGGGCCACGAGACAGTGC 925
Db 721 GGAGAGCAGAGGCTCAGTAAGAGAAATCAAGCATAGTCTGGGCCACGAGACAGTGC 780
Qy 926 TCCACTGGCAGCGCATTTGCTAAATATAGG 955
Db 781 TCCACTGGCAGCGCATTTGCTAAATATAGG 810

RESULT 17
AL565248/c 882 bp mRNA linear EST 05-APR-2004
LOCUS
DEFINITION
AL565248 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DP002YP13 3-PRIME, mRNA sequence.
ACCESSION
AL565248.3 GI:46232132
VERSION
AL565248.3
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 882)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:30549100.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 5584.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DP002CH07NP1&c=5584.f.
Location/Qualifiers
1..882
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DP002YP13"

FEATURES
source

/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 38.5%; Score 800; DB 1; Length 882;
Best Local Similarity 94.2%; Pred. No. 1.2e-195;
Matches 821; Conservative 15; Mismatches 35; Indels 1; Gaps 1;
Qy 1204 AGTGAAGCACTCAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATAC 1263
Db 878 ARGTTGGAGMATTAAAAAAGGTGAATATATGTTTAAAGAGATTAAAGGGCTTTTAM 819
Qy 1264 AAAAATCAGGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAGTGTGAGGCCCTTGA 1323
Db 818 AAAAAAAGAGAGATGAGATATTGAAGAAGAGAAATTAAGAAAGTGTGAAGAAATTTGA 759
Qy 1324 AAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAGCAAGAGCTTGTATTATATAA 1383
Db 758 AAGGAAAAATTTGMAATTTAATAAGAGTTGAATGNNAAAGNAGCTTGTTTTACTAA 699
Qy 1384 CAAAAAGTTAGAGCTCAAGTGGAAGAGCATGCTATCAGAAATCAAAATGAAACAGGCTAA 1443
Db 698 CAAAAAGTTAGAGCTCAAGTGGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTAA 639
Qy 1444 AACAGAGGATGAAAAAGTCCAAATTAATCTGTCTACAGATGACACACAAAGCTTCTTCA 1503
Db 638 AACAGAGGATGAAAAAGTCCAAATTAATCTGTCTACAGATGACACACAAAGCTTCTTCA 579
Qy 1504 AGAACATAATATGCAATTTGAAGCAATTTGAGGAATCAAGAAAGAGTCAAGAAAGT 1563
Db 578 AGAACATAATATGCAATTTGAAGCAATTTGAGGAATCAAGAAAGAGTCAAGAAAGT 519
Qy 1564 GGACAGGGCAGTCTGAAAGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCTGGC 1623
Db 518 GGACAGGGCAGTCTGAAAGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCTGGC 459
Qy 1624 TTCCAAACAGCTGCAAAATGGAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGACCT 1683
Db 458 TTCCAAACAGCTGCAAAATGGAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGACCT 399
Qy 1684 GGAAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCTGCTGA 1743
Db 398 GGAAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCTGCTGA 339
Qy 1744 AAGAGCAGCAGAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCATTGCGCTGGCAGT 1803
Db 338 AAGAGCAGCAGAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCATTGCGCTGGCAGT 279
Qy 1804 TCTGCTGAAAGAGATGATGCTTTTCAAGACCGAGGAGGAGCTCTCTGATGGAGATGCA 1863
Db 278 TCTGCTGAAAGAGATGATGCTTTTCAAGACCGAGGAGGAGCTCTCTGATGGAGATGCA 219
Qy 1864 GAGTCTGATGGGGCAGAGCAAGTCACTCTGACAGAGGCTTACTGTTTCAAGAGG 1923
Db 218 GAGTCTGATGGGGCAGAGCAAGTCACTCTGACAGGAGGCTTACTGTTTCAAGAGG 159
Qy 1924 AGCTGAGGACAGGACTGGGGCAACAGCGGATATTCGATTCATTCTCTGCCCAAGTG 1983
Db 158 AGCTGAGGACAGGACTGGGGCAACAGCGGATATTCGATTCATTCTCTGCCCAAGTG 99
Qy 1984 TGGAGAGGTTCTCCCTGACATAGACAGCTTACAGATTCACGTGATGATGATCATTTTA 2043
Db 98 TGGAGAGGTTCTCCCTGACATAGACAGCTTACAGATTCACGTGATGATGATCATTTTA 39
Qy 2044 AGTGTGATGATACCTCCCAAAAGTGTG 2075
Db 38 AGTGTGAT-TATCACCTCCCSVAAAGTGTG 8

/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 38.0%; Score 789.2; DB 3; Length 882;
Best Local Similarity 97.5%; Pred. No. 7.3e-193;
Matches 811; Conservative 0; Mismatches 20; Indels 1; Gaps 1;
QY 759 TACAAGCAGAGAGGAGAGCTGTGGGATCGTGTCTGAAGTGCAGCTCAAGCTGAAC 818
Db 1 TACAAGCAGAGAGGAGAGCTGTGGGATCGTGTCTGAAGTGCAGCTCAAGCTGAAC 60
QY 819 CCAGCGCTCTCAGAGAGATCTCTTCTGAATAGGATGGCTGAAGGAGAGCAGAG 878
Db 61 CCAGCGCTCTCAGAGAGATCTCTTCTGAATAGGATGGCTGAAGGAGAGCAGAG 120
QY 879 GGTCAAGTAAAGAAATCAAGCATAGTCTCGGCCACGAGAACAGTCTCCACTGGCACGG 938
Db 121 GGTCAAGTAAAGAAATCAAGCATAGTCTCGGCCACGAGAACAGTCTCCACTGGCACGG 180
QY 939 CATGTCTAAATATAGGAGCAGATCTCCAGATGGGCCAAGAAATTAATCTCGAACATGAGG 998
Db 181 CATGTCTAAATATAGGAGCAGATCTCCAGATGGGCCAAGAAATTAATCTCGAACATGAGG 240
QY 999 AGTTAACTGTGAGCCAGCTCTCTGTGCTTGAAGGAGGATCAAGAGTGGAGAGAC 1058
Db 241 AGTTAACTGTGAGCCAGCTCTCTGTGCTTGAAGGAGGATCAAGAGTGGAGAGAC 300
QY 1059 TTCAAGTTGACCTCAAGGAGCCAAAGAAAGATTTTCAGATTTTGAAGAAAGAAAGTA 1118
Db 301 TTGAAGTTGACCTCAAGGAGCCAAAGAAAGATTTTCAGATTTTGAAGAAAGAAAGTA 360
QY 1119 ATCGTTCTGAGATTGAACCCAGACAGAGGGGAGCAGAGAAAGAAATGATGAAGAGA 1178
Db 361 ATCGTTCTGAGATTGAACCCAGACAGAGGGGAGCAGAGAAAGAAATGATGAAGAGA 420
QY 1179 AAGCCCGGAGACTTTGGAAGCGAAGTGAAGCAGTGAACCTCCAGGTGACATCTCTGT 1238
Db 421 AAGCCCGGAGACTTTGGAAGCGAAGTGAAGCAGTGAACCTCCAGGTGACATCTCTGT 480
QY 1239 TTAAGGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGCTAATCAAGAGAGAC 1298
Db 481 TTAAGGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGCTAATCAAGAGAGAC 540
QY 1299 TTCAAGAAAGTGTCAAGGCCCTTTGAAAGGAAAAATTTCTGCAATTTCCATCAGAGTTGAATG 1358
Db 541 TTCAAGAAAGTGTCAAGGCCCTTTGAAAGGAAAAATTTCTGCAATTTCCATCAGAGTTGAATG 600
QY 1359 AAAAGCAGAGCTTTTATATCTAAACAAAGTTAGAGCTTACAGTGGAAAGCATGCTAT 1418
Db 601 AAAAGCAGAGCTTTTATATCTAAACAAAGTTAGAGCTTACAGTGGAAAGCATGCTAT 660
QY 1419 CAGAAATCAAAATGGAACAGGCTTAAACAGAGATGAAAGTCCAAATTAATCTGCTAC 1478
Db 661 CAGAAATCAAAATGGAACAGGCTTAAACAGAGATGAAAGTCCAAATTAATCTGCTAC 720
QY 1479 AGATGACACACAAGCTTTCTTCAAGAACATTAATATGATTTGAAACAAATTTGAGG-AA 1537
Db 721 AGATGACACACAAGCTTTCTTCAAGAACATTAATATGATTTGAAACAAATTTGAGGAAA 780
QY 1538 CTAACAGAAAGAGTCAAGAAAGTGAACAGGCGAGTGTCTGAAGGAGCTGA 1589
Db 781 CTAACAGAAAGAGTCCGAANAAGTGGGACAGGGCGAGTGGCTGAAAGGGA 832

RESULT 20

AL543413

LOCUS

968 bp mRNA linear EST 24-MAR-2004

DEFINITION

AL543413 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI003YG24 5-PRIME, mRNA sequence.

ACCESSION

AL543413

VERSION

AL543413.3 GI:45718973

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 968)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 15, 2001 this sequence version replaced gi:31265260.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

5584.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CSODI003BD12QP1&c=5584.f.

Location/Qualifiers

1..968

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI003YG24"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 37.4%; Score 777.2; DB 1; Length 968;

Best Local Similarity 94.4%; Pred. No. 9.7e-190;

Matches 829; Conservative 11; Mismatches 18; Indels 20; Gaps 3;

QY 297 CAGGAATCTCTCAATGTCCCATCCATCAACCTCTCASCCTCACTKAAAAGGAGGACAGCC 356

Db 104 CAGGAATCTCTCAATGTCCCATCCATCAACCTCTCASCCTCACTKAAAAGGAGGACAGCC 163

QY 357 CCAGTGAAGACAGAGAAATGGACCCGCCACCTGGCCCAACCCCAACCTGGACACGTTTA 416

Db 164 CCAGTGAAGACAGAGAAATGGACCCGCCACCTGGCCCAACCCCAACCTGGACACGTTTA 223

QY 417 CCCCAGGAGGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAAACCCAGCTGAAAG 476

Db 224 CCCCAGGAGGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAAACCCAGCTGAAAG 283

QY 477 AAGCCATGAGCTTAATATCAAGCCATGAAGGGGAGATTGAGGAGCTTTCGGCTGGA 536

Db 284 AAGCCATGAGCTTAATATCAAGCCATGAAGGGGAGATTGAGGAGCTTTCGGCTGGA 343

QY 537 CAGAGAAAACAGAGGAGAAACGCCAGTCTTTTGTAGATACAGAGCAAGAGCAAGCAAGAGAGC 596

Db 344 CAGAGAAAACAGAGGAGAAACGCCAGTCTTTTGTAGATACAGAGCAAGAGCAAGCAAGAGAGC 403

QY 597 GTCTAATGCGCTTGAGTCTATGAGAATGAGAAATTTGAAGGAGAGCTTTGAGAACTAAAAG 656

Db 404 GTCTAATGCGCTTGAGTCTATGAGAATGAGAAATTTGAAGGAGAGCTTTGAGAACTAAAAG 463

QY 657 GGAAATCAGAAAGGTCTCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAG 716

Db 464 GGAAATCAGAAAGGTCTCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAG 523

Qy	717	CGGAGCAGGAAAGGACCAAGCTCAGACCCAGGTCGAGGCTCAACAGCAGAGAGGCAG	776
Db	524	CCGAGCAGGAAAGGACCAAGCTCAGACCCAGGTCGAGGCTCAACAGCAGAGAGGCAG	583
Qy	777	ACCTGTTGGGCATCGTGTCTGAAGTGGCTGAAGGAGAGCAGAGGGTCCCTCAGAG	836
Db	584	ACCTGTTGGGCATCGTGTCTGAAGTGGCTGAAGGAGAGCAGAGGGTCCCTCAGAG	643
Qy	837	ATTCTCTTTGTTGAAATTTAGGATGGCTGAAGGAGAGCAGAGGGTCACTAAGAAATCA	896
Db	644	ATTCTCTTTGTTGAAATTTAGGATGGCTGAAGGAGAGCAGAGGGTCACTAAGAAATCA	703
Qy	897	AGCATAGTCTCTGGGCCCCAGAGAACAGTCTCCACTGGCAGCGCATTTGCTTAAATATAGGA	956
Db	704	AGCATAGTCTCTGGGCCCCAGAGAACAGTCTCCACTGGCAC-----GA	745
Qy	957	GCAGATCTCAGATAGGGGCCAAGAAATTACTTTCGAACATGAGAGTAACTGTGAGGCCAGC	1016
Db	746	GCAGATCTCAGATAGGGGCCAAGAAATTACTTTCGAACATGAGAGTAACTGTGAGGCCAGC	805
Qy	1017	TCCTGCTGTCCTTAAGGAGGGAATCAGAAGGTGGAGAGCTTCAAGTTCGCATCTCAAGG	1076
Db	806	TCCTGCTGTCCTTAAGGAGGGAATCAGAAGGTGGAGAGCTTCAAGTTCGCATCTCAAGG	865
Qy	1077	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTCAGATTTGAAA	1136
Db	866	RGGCCAAAGAAAGAGTTTCAGATTTTGAAG-AAAAG-AAAAGTAATCGTTCTCAGATTTAAC	924
Qy	1137	CCGAGCAGAGGGGAGCAGAGAAAGAGAAATGATGAA	1174
Db	925	CAGACAGAGGGGSA-CAGAGAAAGATKATGAAAAA	961
RESULT 21			
LOCUS	CN309549	772 bp	mRNA linear EST 16-MAY-2004
DEFINITION	17000599888206 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.		
ACCESSION	CN309549		
VERSION	CN309549.1	GI:47325963	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 772) Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J. and Stanton, L. W.		
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation		
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)		
PUBMED	15146197		
COMMENT	Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 772 Std Error: 0.00.		
FEATURES	Location/Qualifiers		
source	1..772		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="embryonic stem cell, retinoic acid and		
	mitogen-treated hES cell line H7"		
	/clone_lib="GRN_PRENEU"		
	/notes="oligo dr primed, full-length enriched cDNA library		
	from hES cell line H7 (p29) maintained in feeder-free		
	conditions. Embryoid bodies were generated in the presence		
	of all-trans retinoic acid and mitogens."		
	1 (bases 1 to 817)		
ORIGIN			
Query Match 37.2%; Score 772; DB 8; Length 772;			
Best Local Similarity 100.0%; Pred. No. 2e-188;			
Matches 772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	938	GCATTGTCTAAATATAGGAGCAGATCTGTCAGATGGGCGCAAGAAATTACTTCGAACATGAG	997
Db	1	GCATTGTCTAAATATAGGAGCAGATCTGTCAGATGGGCGCAAGAAATTACTTCGAACATGAG	60
Qy	998	GAGTTAACTGTGAGCCAGCTCTGCTGTGTCCTTAAGGGAAGGAATCAGAAGGTGGAGAGA	1057
Db	61	GAGTTAACTGTGAGCCAGCTCTGCTGTGTCCTTAAGGGAAGGAATCAGAAGGTGGAGAGA	120
Qy	1058	CTTGAAGTTGCACTCAAGSGGSCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAGT	1117
Db	121	CTTGAAGTTGCACTCAAGSGGSCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAGT	180
Qy	1118	AATCGTTCTGAGATTGAAACCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAG	1177
Db	181	AATCGTTCTGAGATTGAAACCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAG	240
Qy	1178	AAAGCCCCGAGACATGTTTGAAGCGAAGTGAAGACATGAACCTCCAGGTGACATCTCTG	1237
Db	241	AAAGCCCCGAGACATGTTTGAAGCGAAGTGAAGACATGAACCTCCAGGTGACATCTCTG	300
Qy	1238	TTTAAGGAGCTTCAGAGGCTCATACAAACTCAGCGAAGCTAGCTAATCAAGAGAGA	1297
Db	301	TTTAAGGAGCTTCAGAGGCTCATACAAACTCAGCGAAGCTAGCTAATCAAGAGAGA	360
Qy	1298	CTTCAAGAAAAGTGTCAAGGCCCTTGAAGGAAAATTCTGCAATTCATCAGAGTTGAAT	1357
Db	361	CTTCAAGAAAAGTGTCAAGGCCCTTGAAGGAAAATTCTGCAATTCATCAGAGTTGAAT	420
Qy	1358	GAAAAGCAAGAGCTGTTTATATCAACAAAAAGTTAGAGCTACAAGTGGAAAAGCATGCTA	1417
Db	421	GAAAAGCAAGAGCTGTTTATATCAACAAAAAGTTAGAGCTACAAGTGGAAAAGCATGCTA	480
Qy	1418	TCAGAAATCAAAATGGAACAGCTTAAACACAGAGATGAAAAGTCCAAATTAAGTGTGCTA	1477
Db	481	TCAGAAATCAAAATGGAACAGCTTAAACACAGAGATGAAAAGTCCAAATTAAGTGTGCTA	540
Qy	1478	CAGATGACACAAACAAGCTCTTCAAGAACATAATATGATTCATGAAAAACAATTGAGGAA	1537
Db	541	CAGATGACACACAAACAAGCTCTTCAAGAACATAATATGATTCATGAAAAACAATTGAGGAA	600
Qy	1538	CTAACAGAAAAAGAGCTCAGAAAAAGTGGACAGGCGAGCTGTGGAAGGAACGTAGTGAAAAA	1597
Db	601	CTAACAGAAAAAGAGCTCAGAAAAAGTGGACAGGCGAGCTGTGGAAGGAACGTAGTGAAAAA	660
Qy	1598	CTGGAAGTGGCAGAGAGGCTCTGGCTTCCAAAACAGCTGCAAAATGGATGAATCAAGCAA	1657
Db	661	CTGGAAGTGGCAGAGAGGCTCTGGCTTCCAAAACAGCTGCAAAATGGATGAATCAAGCAA	720
Qy	1658	ACCATTTGCCAAGCAGAGAGGACCTCGAAAACCATGACCATCTCCAGGGCTC	1709
Db	721	ACCATTTGCCAAGCAGAGAGGACCTCGAAAACCATGACCATCTCCAGGGCTC	772
RESULT 22			
LOCUS	BU623283/c	817 bp	mRNA linear EST 23-SEP-2002
DEFINITION	UI-H-FLI-bgc-g-19-0-UI-s1 NCI_CGAP_Ful Homo sapiens cDNA clone		
ACCESSION	BU623283		
VERSION	BU623283.1	GI:23289498	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 817)		

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Martin
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Location/Qualifiers
1. .817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FLI-bgc-g-19-0-UI"
/tissue_types="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_FLI"
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP_FLI is a normalized CDNA library
derived from a pool of mRNA obtained from 4 cell lines
from grade III chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
GAGTCCGGTG. The cell lines were provided by Dr. James
Martin from the University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FLI
TAG_SEQ=GAGGTCGGTG"

ORIGIN

Query Match 37.1%; Score 771.4; DB 3; Length 817;
Best Local Similarity 99.6%; Pred. No. 2.9e-188;
Matches 772; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1303 AGAAAAAGTGTGAGGCGCTTGAAGGAAAAATCTGCAATTCATCAGAGTTGAATGAAAA 1362
DB 817 AGAAAAAGTGTGAGGCGCTTGAAGGAAAAATCTGCAATTCATCAGAGTTGAATGAAAA 758
QY 1363 GCAAGAGCTGTTTATCTAACAAAAAGTTAGAGCTACAGTGAAGGAAAGCATGTATCAGA 1422
DB 757 GCAAGAGCTGTTTATCTAACAAAAAGTTAGAGCTACAGTGAAGGAAAGCATGTATCAGA 698
QY 1423 AATCAAAATGGAACAGCGCTTAACACAGGATGAAGTCCAAATTAAGTGTGCTACAGAT 1482
DB 697 AATCAAAATGGAACAGCGCTTAACACAGGATGAAGTCCAAATTAAGTGTGCTACAGAT 638
QY 1483 GACACACAAAGCTTCTTCAAGAACATATAATGCAATGAAACAAATGAGGAACTAAC 1542
DB 637 GACACACAAAGCTTCTTCAAGAACATATAATGCAATGAAACAAATGAGGAACTAAC 578
QY 1543 AGAAAAAGAGTCAGAAAAAGTGACAGGGCAGTGTCTGAAGGAAGCTGAGTGAAGAACTGGA 1602
DB 577 AGAAAAAGAGTCAGAAAAAGTGACAGGGCAGTGTCTGAAGGAAGCTGAGTGAAGAACTGGA 518
QY 1603 ACTGGCAGAGAGGCTCTCGCTTCCAAACAGCTGCAANTGGATGAATGAAGCAACCAT 1662
DB 517 ACTGGCAGAGAGGCTCTCGCTTCCAAACAGCTGCAANTGGATGAATGAAGCAACCAT 458

QY 1663 TGCCAAAGCAGGAGAGGAGCACTGGAACCATGACCATCTCCTCAGGCTCAGATGGAAGTTTA 1722
DB 457 TGCCAAAGCAGGAGAGGAGCACTGGAACCATGACCATCTCCTCAGGCTCAGATGGAAGTTTA 398
QY 1723 CTGTTCTGATTTTCATGCTGAAAGAGCAGGAGAGAGAAAAATTCATGAGGAAAGGAGCA 1782
DB 397 CTGTTCTGATTTTCATGCTGAAAGAGCAGGAGAGAGAAAAATTCATGAGGAAAGGAGCA 338
QY 1783 ACTGGCATTGCAGCTGGCAGTCTGCTGAAAGAGAAATGATGCTTTCGAACACGGAGCAG 1842
DB 337 ACTGGCATTGCAGCTGGCAGTCTGCTGAAAGAGAAATGATGCTTTCGAACACGGAGCAG 278
QY 1843 GCAGTCTTGTATGAGATCCAGAGTCGTATGATGGGGGAGAGCAAGTGAATCTCTGACGAGCA 1902
DB 277 GCAGTCTTGTATGAGATCCAGAGTCGTATGATGGGGGAGAGCAAGTGAATCTCTGACGAGCA 218
QY 1903 GGCCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGGAGCTGGCGCAACAGCGGAATATTCC 1962
DB 217 GGCCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGGAGCTGGCGCAACAGCGGAATATTCC 158
QY 1963 GATTCAATCTCTGCCCCCAAGTGTGGAGAGGTTCGCTGACATAGACACGTTACAGATTCA 2022
DB 157 GATTCAATCTCTGCCCCCAAGTGTGGAGAGGTTCGCTGACATAGACACGTTACAGATTCA 98
QY 2023 CGTGATGGATTGCATCATTTAAAGTGTGATGATATCACCTCCCAAAACTGTTGGT 2077
DB 97 CGTGATGGATTGCATCATTTAAAGTGTGATGATATCACCTCCCAAAACTGTTGGT 43
RESULT 23
BE875862 945 bp mRNA linear EST 20-OCT-2000
LOCUS 601487337F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3889542 5',
DEFINITION mRNA sequence.
ACCESSION BE875862
VERSION BE875862.1 GI:10324638
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 945)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9671 row: f column: 07
High quality sequence stop: 675.
Location/Qualifiers
1. .945
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3889542"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 69"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."

Query Match 37.1%; Score 770; DB 7; Length 945;

Best Local Similarity 96.5%; Pred. No. 7.1e-188;		Matches 808; Conservative 0; Mismatches 26; Indels 3; Gaps 2;	
QY	565	TTTTGAGATACAGAGCAAGAGCAAAAGAGCGTCTAATGGCCTTGAGTCATGAGAATGA	624
Db	1	TTTTGAGATACAGAGCAAGAGCAAAAGAGCGTCTAATGGCCTTGAGTCATGAGAATGA	60
QY	625	GAATTTGAAGGAGAGCTTGGAAAACTAAAGGGAAATCAGAAAAAGTCTATCGAGACCC	684
Db	61	GAATTTGAAGGAGAGCTTGGAAAACTAAAGGGAAATCAGAAAAAGTCTATCGAGACCC	120
QY	685	CACTGATGATCTCCAGGCTTCCAGGCGCGAAGCGGAGCAGGAGGAAAAAGCACCTCAGGAC	744
Db	121	CACTGATGATCTCCAGGCTTCCAGGCGCGAAGCGGAGCAGGAGGAAAAAGCACCTCAGGAC	180
QY	745	CCAGGTGGTGAAGCTACAGCAGAGAGAGGACAGACCTGTTGGGCATCGTCTGAACTGCA	804
Db	181	CCAGGTGGTGAAGCTACAGCAGAGAGAGGACAGACCTGTTGGGCATCGTCTGAACTGCA	240
QY	805	GCTCAAGCTGAATCTCCAGCGGCTCCTCAGAAGATTCTTTGTTGAAATTTAGGATGCTGA	864
Db	241	GCTCAAGCTGAATCTCCAGCGGCTCCTCAGAAGATTCTTTGTTGAAATTTAGGATGCTGA	300
QY	865	AGGAGAGCAGAGGCTAGTAAAAAGAAATCAAGCATAGTCTCTGGGCCACAGAACAGT	924
Db	301	AGGAGAGCAGAGGCTAGTAAAAAGAAATCAAGCATAGTCTCTGGGCCACAGAACAGT	360
QY	925	CTCCACTGCACGGCATTGCTTAATATATAGGAGCAGATCTGAGATGGGCCCAAGATTA	984
Db	361	CTCCACTGCACGGCATTGCTTAATATATAGGAGCAGATCTGAGATGGGCCCAAGATTA	420
QY	985	CTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCTCTGCTGCTGCTAAAGGAGGAATCA	1044
Db	421	CTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCTCTGCTGCTGCTAAAGGAGGAATCA	480
QY	1045	GAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAAGAAAGAGTTTCAGATTTTGA	1104
Db	481	GAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAAGAAAGAGTTTCAGATTTTGA	540
QY	1105	AAAGAAAAAAGTAATCGTTCTGAGATTGAATCCAGCAGAGAGGGAGCAGAGAAAGA	1164
Db	541	AAAGAAAAAAGTAATCGTTCTGAGATTGAATCCAGCAGAGAGGGAGCAGAGAAAGA	600
QY	1165	GAATGATGAAGAGAAAGGCCGAGACTGTTGGAAGCGAAGTGGAGCACTCAACTCCA	1224
Db	601	GAATGATGAAGAGAAAGGCCGAGACTGTTGGAAGCGAAGTGGAGCACTCAACTCCA	659
QY	1225	GGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGT	1284
Db	660	NGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGTTGAGTA	719
QY	1285	AATGAAGAAGACTTCAAGAAAAAGTGTGAGGCCCTTGAAGGAAAAAATTCGCAATTC	1344
Db	720	ATGGAAGAAGAGACTTCAAGAAAAAGTGTGAGGCCCTTGAAGGAAAAAATTCGCAATTC	777
QY	1345	ATCAGAGTTGAATGAAAAAGCAAGAGCTCTTTTATACACAAAAAGTTAGAGCTACA	1401
Db	778	ATCAGAGTTGAATGAAAAAGCAAGAGCTTGGTTATACAAAAAGTTGAGGCTACA	834

RESULT 24
CD243803
LOCUS
DEFINITION
AGENCOURT_14099182 NIH_MGC_180 Homo sapiens CDNA clone
IMAGE:30377971 5', mRNA sequence.
ACCESSION
CD243803
VERSION
CD243803.1 GI:31004267
KEYWORDS
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE	1 (bases 1 to 854)	NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)	
TITLE	Unpublished (1999)	
JOURNAL	Contact: Robert Strausberg, Ph.D.	
COMMENT	Email: cgabbs-remail.nih.gov	
	Tissue Procurement: Dr. Michael Brownstein	
	cDNA Library Preparation: Invitrogen Corp	
	cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)	
	Clone distribution: MGC clone distribution information can be	
	found through the I.M.A.G.E. Consortium/LLNL at:	
	http://image.llnl.gov	
	Plate: NDAM443 row: a column: 20	
	High quality sequence stop: 615.	
FEATURES	location/Qualifiers	
source	1..854	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clones="IMAGE:30377971"	
	/lab_host="DH10B-ron A (T1 and T5 phage resistances)"	
	/clone_lib="NIH_MGC_180"	
	/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;	
	Site 2: EcoRV (destroyed); Library is oligo-dT primed and	
	directionally cloned (EcoRV site is destroyed upon	
	cloning). Average insert size 1.68 kb. Library was	
	constructed by (Invitrogen). Note: this is a NIH_MGC	
	Library."	
ORIGIN	Query Match 36.8%; Score 763.8; DB 5; Length 854;	
	Best Local Similarity 99.7%; Pred. No. 2.8e-186;	
	Matches 765; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	421	GGAGGAGCTGCTGCAGCAGATGAAAGAGCTCTGTACCAAGAACCCACAGCTGAAAGAGC 480
Db	1	GGAGGAGCTGCTGCAGCAGATGAAAGAGCTCTGTACCGAGAACCCACAGCTGAAAGAGC 60
QY	481	CATGAAGCTAAATATCAAGCCNTGAAAGGAGATTTGAGGAGCTTTGCGCTTGACAGAGA 540
Db	61	CATGAAGCTAAATATCAAGCCCATGAAAGGAGATTTGAGGAGCTTTGCGCTTGACAGAGA 120
QY	541	GAACAGAGAGGAAGAACCCAGTCTTTTGTAGATACAGAGCAAGAAAGCAAGAGAGCTCT 600
Db	121	GAACAGAGAGGAAGAACCCAGTCTTTTGTAGATACAGAGCAAGAAAGCAAGAGAGCTCT 180
QY	601	AATGCGCTTGAGTCATGAGAAATGAGAAATTGAAAGGAAGAGCTTGGAATACTAAAAAGG 660
Db	181	AATGCGCTTGAGTCATGAGAAATGAGAAATTGAAAGGAAGAGCTTGGAATACTAAAAAGG 240
QY	661	ATCAGAAAGGTCATCTGAGGAGCCCACTGATGACTCCAGGCTTCCAGGGCCGAGAGCGGA 720
Db	241	ATCAGAAAGGTCATCTGAGGAGCCCACTGATGACTCCAGGCTTCCAGGGCCGAGAGCGGA 300
QY	721	CGAGAAAGGACAGAGCTCAGAGCCAGTGTGAGGCTTACAAAGCAGAGAGAGCGAGACT 780
Db	301	CGAGAAAGGACAGAGCTCAGAGCCAGTGTGAGGCTTACAAAGCAGAGAGAGCGAGACT 360
QY	781	GTTGGGCATCGTGTGTAACCTGCAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAGATTC 840
Db	361	GTTGGGCATCGTGTGTAACCTGCAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAGATTC 420
QY	841	CTTTGTTGAAATTTAGGATGGCTGAGGAGAGCAGAGAGGTCAGTAAAGAAATCAAGCA 900
Db	421	CTTTGTTGAAATTTAGGATGGCTGAGGAGAGCAGAGAGGTCAGTAAAGAAATCAAGCA 480
QY	901	TAGTCTGGGCCCCAGAGAACAGTCTCCACTGGCACCGCATTTGTCTAAATATAGGAGCAG 960
Db	481	TAGTCTGGGCCCCAGAGAACAGTCTCCACTGGCACCGCATTTGTCTAAATATAGGAGCAG 540
QY	961	ATCTGAGATCGGGCCAAAGAAATTAATCTCGAAATGAGGAGTTAACTGTGAGCAGCTCT 1020


```

Db      541 ATCTGCGATGGGGCCAAAGAAATTAATTCTTCGAACATGAGGAGTTAACTCTGAGCCAGCTCCT 600
Qy      1021 GCTGTGCTTAAGGGAAGGGGAATCAGAGGTGAGAGACTTTGAAGTTGCACTCAAGGAGGC 1080
Db      601 GCTGTGCTTAAGGGAAGGGGATCAGAGGTGAGAGACTTTGAAGTTGCACTCAAGGAGGC 660
Qy      1081 CAAAGAAGAGTTTCAGATTTTGAAGAAGAAAAACAAGTAATCGTTCTGAGATTGAAACCCA 1140
Db      661 CAAAGAAGAGTTTCAGATTTTGAAGAAGAAAAACAAGTAATCGTTCTGAGATTGAAACCCA 720
Qy      1141 GACAGAGGGGAGCACAGAGAAAGAGATGATGAAGAGAAAGGCCCG 1187
Db      721 GACAGAGGGGAGCACAGAGAAAGAGATGATGAAGAGAAAGGCCCG 767

RESULT 25
BX364738
LOCUS
DEFINITION
BX364738 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Accession
BX364738
Version
BX364738.2 GI:46286045
Keywords
EST.
Source
Homo sapiens (human)
Organism
Homo sapiens
Reference
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
Authors
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Title
Full-length cDNA libraries and normalization
Journal
Unpublished (2001)
Comment
On May 5, 2003 this sequence version replaced gi:30366842.
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5584.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS1A10032B100P1k=5584.f.
FEATURES
source
1..990
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI009YB18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 36.0%; Score 747.6; DB 4; Length 990;
Best Local Similarity 90.8%; Pred. No. 4.6e-182;
Matches 799; Conservative 30; Mismatches 29; Indels 22; Gaps 3;
Qy      297 CAGGAATCTTCGAATGCTCCATCACTCTCAGCTGCCTCACTGAAAGGAGGAGCC 356
Db      131 CAGGAATCTTCGAATGCTCCATCACTCTCAGCTGCCTCACTGAAAGGAGGAGCC 190
Qy      357 CCAGTGAAGCAGAGAAATGGACCCGCCACCTGGCCACCCAAACCTGGACACCTTTA 416
Db      191 CCAGTGAAGCAGAGAAATGGACCCGCCACCTGGCCACCCAAACCTGGACACCTTTA 250
Qy      417 CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAACCCAGCTGAAAG 476

```

```

Db      251 CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCCTGACCGAGAACCCAGCTGAAAG 310
Qy      477 AAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTTTGAAGGAGCTTTCCGCTCGGA 536
Db      311 AAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTTGAGGAGCTTTCCGCTCGGA 370
Qy      537 CAGAGAAA CAGAGAAAGAACCGCCAGCTTTTGTGAGATACAGAGCAAAAGCAAAAGAGC 596
Db      371 CAGAGAAA CAGAGAAAGAACCGCCAGCTTTTGTGAGATACAGAGCAAAAGCAAAAGAGC 430
Qy      597 GTCTAATGCGCTTGAGTCATGAGAAATGAGAAATGAGAAAGAGCTTGGAAAACTAAAAG 656
Db      431 GTCTAATGCGCTTGAGTCATGAGAAATGAGAAATGAGAAAGAGCTTGGAAAACTAAAAG 490
Qy      657 GGAATATCAGAAAGGTCTATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGGAAG 716
Db      491 GGAATATCAGAAAGGTCTATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGGAAG 550
Qy      717 CGGAGCAGGAAAAGGACCAAGCTCAGGACCCAGGTGTGAGGCTTCAAGCAGAGAAAGGCGAG 776
Db      551 CGGACAGTGAAGAAAGGACCAAGCTCAGGACCCAGGTGTGAGGCTTCAAGCAGAGAAAGGCGAG 610
Qy      777 ACTGTGTGGGCATCGTCTGAACTGCACTCAAGCTCAAGCTCCAGGGCTCCTCAGAAG 836
Db      611 ACTGTGTGGGCATCGTCTGAACTGCACTCAAGCTCAAGCTCCAGGGCTCCTCAGAAG 670
Qy      837 ATTCTCTTGTGAAATTAGGATGGCTGAAGGAGAACAGAGGTTCAGTAAAAAGAAATCA 896
Db      671 ATTCTCTTGTGAAATTAGGATGGCTGAAGGAGAACAGAGGTTCAGTAAAAAGAAATCA 730
Qy      897 AGCATAGTCTCTGGGCCACGAGAACAGTCTCCACTCGGCAGCGCATTTGTCTAAATATAGGA 956
Db      731 AGCATAGTCTCTGGGCCACGAGAACAGTCTCCACTCGGCAC-----GA 772
Qy      957 GCAGATCTCAGATGGGGCCAAAGAAATTAATTCTGAAACATGAGAGTTAACTGTGAGCCAGC 1016
Db      773 GCAGATCTCAGATGGGGCCAAAGAAATTAATTCTGAAACATGAGAGTTAACTGTGAGCCAGC 832
Qy      1017 TCTGTGTGCTTAAGGGGAAGGGAATCAGAGGTGAGAGACTTGAAGTTGCACTCAAGG 1076
Db      833 YCTGTGTGCTTAAGGGGAAGGGAATCAGAGGTGAGAGACTTGAAGTTGCACTCAAGG 892
Qy      1077 AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAATCGTTCTGAGATTGAAA 1136
Db      893 AGSCAAAGAAAGAGTTTACG---TTTGAAAAAGAAAMWARKATCSKKCKGAGATKG-AA 948
Qy      1137 CCCAGACAGAGGGGACACAGAGAAAGAGAAATGATGAAGA 1176
Db      949 CCCARACASAGGGGRSCMARAGAAAAAARWGAAGARAA 988

RESULT 26
AL565831/c
LOCUS
DEFINITION
AL565831 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF008YE13 3-PRIME, mRNA sequence.
Accession
AL565831
Version
AL565831.3 GI:46232652
Keywords
EST.
Source
Homo sapiens (human)
Organism
Homo sapiens
Reference
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
Authors
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Title
Full-length cDNA libraries and normalization
Journal
Unpublished (2001)
Comment
On Feb 16, 2001 this sequence version replaced gi:30550229.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr

```


RESULT 29
AL538121
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FETAL BRAIN Homo sapiens FETAL BRAIN Homo sapiens CDNA clone
5-PRIME, mRNA sequence.
AL538121
AL538121
AL538121.3 GI:45713865
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 989)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:31262722.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 5584.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODF029CH03QP1&c=5584.f.
Location/Qualifiers
1..989
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF029YP05"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 35.1%; Score 728; DB 1; Length 989;
Best Local Similarity 97.4%; Pred. No. 5.6e-177;
Matches 773; Conservative 8; Mismatches 8; Indels 5; Gaps 4;
QY 297 CAGGAACTTCTGCAATGTCCCATCACTCTAGCTGCTCACTGAAAGGAGGACGCC 356
DB 199 CAGGAACTTCTGCAATGTCCCATCACTCTAGCTGCTCACTGAAAGGAGGACGCC 258
QY 357 CAGTGAAGACAGGAAATGACCCGCCACCTGGCCACCAACCTGGACACCTTTA 416
DB 259 CAGTGAAGACAGGAAATGACCCGCCACCTGGCCACCAACCTGGACACATTTA 318
QY 417 CCCCGAGGAGCTGTGTCAGCAGATGAAGAGCTCTGACCAAGAACCAACGCTGAAAG 476
DB 319 CCCCGAGGAGCTGTGTCAGCAGATGAAGAGCTCTGACCCAGAACCAACGCTGAAAG 378
QY 477 AACCCATGAAGCTAAATATCAAGCATGAAGGGAGATTTCAGGAGCTTCGGCTGGA 536
DB 379 AAGCCATGAAGCTAAATATCAAGCATGAAGGGAGATTTCAGGAGCTTCGGCTGGA 438
QY 537 CAGAGAAAACAGAGGAAAGACGCCAGCTTTTTCAGATACAGAGCAAGAAAGCAAGAGGC 596
DB 439 CAGAGAAAACAGAGGAAAGACGCCAGCTTTTTCAGATACAGAGCAAGAAAGCAAGAGGC 498
QY 597 GTCTAATGCGCTTGTAGTCTAGATGAGAAATTTGAGGAAAGAGCTTGAAAACTAAAG 656

Db 499 GTCTAATGCGCTTGTAGTCTAGATGAGAAATTTGAGGAAAGAGCTTGAAAACTAAAG 558
QY 657 GGAATCAGAAAGTCTATCTGAGGACCCCACTGATGATCCAGGCTTCCAGGGCCGAG 716
Db 559 GGAATCAGAAAGTCTATCTGAGGACCCCACTGATGATCCAGGCTTCCAGGGCCGAG 618
QY 717 CGGAGCAGAAAAAGGACCCAGCTCAGGACCCAGGCTGAGGCTCAAGCAGAGAAAGCCAG 776
Db 619 CGGA-SAGAAAAAGACCCAGCTCAGAACCCAGGCTGAGGCTACAGCAGAGAAAGCCAG 677
QY 777 ACCTGTTGGGCATCGTGTCTGAATGTCAGCTCAAGCTCAAGCTCCAGCGCTCTCCAGAG 836
Db 678 ACCTGTTGGGCATCGTGTCTGAATGTCAGCTCAAGCTCAAGCTCCAGCGCTCTCCAGAG 737
QY 837 ATTCCTTTTGTGAATTTAGGATGGCTGAGGAGAGCAGAGGGTCACTTAAAGAAATCA 896
Db 738 ATTCCTTTTGTGAATTTAGGATGGCTGAGGAGAGCAGAGGGTCACTTAAAGAAATCA 797
QY 897 AGCATAGTCTCTGGGCCACGAGAACAGTCTCCACTGGCACCGCATTTGCTAAATATAGGA 956
Db 798 AGCATAGTCTCTGGGCCACGAGAACAGTCTCCACTGGCACCGCATTTGCTAAATATAGGA 857
QY 957 CGAGATCTGCAGAT-GGGGCCAAGAAATTAATTTCGAACATGAGAGATTAACTGTGAGCCAG 1015
Db 858 GCAGATCTGCAGATGGGGGCCAAGAAATTAATTTCGAACATGAGAGATTAACTGTGAGCCAR 917
QY 1016 CTCCTGCTGTCCTTAAGGAGGAGGAATCAGAGAGTGGAGAGACTTGAAGTTGCATCAAG 1075
Db 918 CTCCTGCTGTCCTTAAGGAGGAGGG-ATCARAAGGTGGAGAGAYT--GAAGTTSMCTCAAG 974
QY 1076 GAGGCCAAAGAAAG 1089
Db 975 GAGGCCAAAGAAAG 988
RESULT 30
BQ883124
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BQ883124
BQ883124.1 GI:22275132
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13566 row: e column: 06
High quality sequence stop: 690.
Location/Qualifiers
1..872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6181613"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"

/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACAGCGTGG-3' and
5'-GACTAGTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match 34.9%; Score 725.6; DB 3; Length 872;
Best Local Similarity 99.1%; Pred. No. 2.2e-176;
Matches 739; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 297 CAGGAACCTTGCATCTCCATCAACCTCTCAGCTGCCTCACTGAAAGGAGGACAGCC 356
Db 111 CAGGAACCTTGCATCTCCATCAACCTCTCAGCTGCCTCACTGAAAGGAGGACAGCC 170

QY 357 CCAGTGAAGACAGAGGAATGGACCCGCCACCTGGCCACCCAAACCTGGACACGTTTA 416
Db 171 CCAGTGAAGACAGAGGAATGGACCCGCCACCTGGCCACCCAAACCTGGACACGTTTA 230

QY 417 CCCGGAGGAGCTGCTGCAGCAGATGAAGAGCTCTTGACCAAGAACCCACAGCTGAAG 476
Db 231 CCCGGAGGAGCTGCTGCAGCAGATGAAGAGCTCTTGACCAAGAACCCACAGCTGAAG 290

QY 477 AAGCCATGAAGCTAAATAATCAAGCCATGAAGAGGAGATTGAGGAGCTTTCGGCTTGA 536
Db 291 AAGCCATGAAGCTAAATAATCAAGCCATGAAGAGGAGATTGAGGAGCTTTCGGCTTGA 350

QY 537 CAGAGAAACAGAGGAAGAGCGCAGTTTTTTCAGATACAGAGCAAGAGCAAGAGCAAGAGC 596
Db 351 CAGAGAAACAGAGGAAGAGCGCAGTTTTTTCAGATACAGAGCAAGAGCAAGAGCAAGAGC 410

QY 597 GTCTAATGGCTTTGAGTCATGAGAAATGAGAAATTTGAAGGAAGAGCTTGGAAAACTTAAAG 656
Db 411 GTCTAATGGCTTTGAGTCATGAGAAATGAGAAATTTGAAGGAAGAGCTTGGAAAACTTAAAG 470

QY 657 GGAATCAGAAAGGTCTCTGAGAGACCCCTCATGATGATCTCAGGCTTCCAGGGCCGAAG 716
Db 471 GGAATCAGAAAGGTCTCTGAGAGACCCCTCATGATGATCTCAGGCTTCCAGGGCCGAAG 530

QY 717 CGGAGCAGGAAAGGACAGCTCAGGACCCAGTGTGAGGCTACAGCAGAGAGGCGAG 776
Db 531 CGGAGCAGGAAAGGACAGCTCAGGACCCAGTGTGAGGCTACAGCAGAGAGGCGAG 590

QY 777 ACCTGTTGGGCATCGTGTCTGAACTGAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAG 836
Db 591 ACCTGTTGGGCATCGTGTGAACTGAGCTCAAGCTCAAGCTCAAGCTCAGCGGCTCTCTCAGAAG 650

QY 837 ATTCCTTTTGTGAAATTTAGGATGGCTGAAGGAGAGCAGAGGCTCAGTAAAAAGAAATCA 896
Db 651 ATTCCTTTTGTGAAATTTAGGATGGCTGAAGGAGAGCAGAGGCTCAGTAAAAAGAAATCA 710

QY 897 AGCATAGTCTGGGCCCAAGAGACAGTCTCCACTGCGCAGGATGCTTAAATATAGGA 956
Db 711 AGCATAGTCTGGGCCCAAGAGACAGTCTCCACTGCGCAGGATGCTTAAATATAGGA 770

QY 957 GCAGATCTGCAGATGGGCCCAAGAAATTACTTTCGAACATGAGGAGTTAACT-GTAGCCAG 1015
Db 771 GCAGATCTGCAGATGGGCCCAAGAAATTACTTTCGAACATGAGGAGTTAACTGTGTGAGCCAG 830

QY 1016 CTCCTGCTGTGCTTAAGGGAGGGAA 1041
Db 831 CTCCTGCTGTGCTTAAGGGAGGGGA 856

RESULT 31

AL535231 907 bp mRNA linear EST 24-MAR-2004
LOCUS

DEFINITION

AL535231 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF008YE13 5-PRIME, mRNA sequence.

ACCESSION

AL535231

VERSION

AL535231.3 GI:457111109

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 907)

AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

On Feb 13, 2001 this sequence version replaced gi:30541800.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Creneau, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

was not normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

This sequence belongs to sequence cluster 5584.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DF008AC07QPl&c=5584.f.

FEATURES

Location/Qualifiers

1..907

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DF008YE13"

/tissue_type="FETAL BRAIN"

/dev_stage="fetal"

/clone_lib="Homo sapiens FETAL BRAIN"

/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."

ORIGIN

Query Match 34.7%; Score 721.4; DB 1; Length 907;

Best Local Similarity 97.3%; Pred. No. 2.8e-175;

Matches 750; Conservative 1; Mismatches 2; Indels 18; Gaps 1;

QY 297 CAGGAACCTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGGAGGACAGCC 356

Db 155 CAGGAACCTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGGAGGACAGCC 214

QY 357 CAGTGAAGACAGAGAAATGGACCCGCCACCTGGCCACCCAAACCTTGAACACGTTTA 416

Db 215 CAGTGAAGACAGAGAAATGGACCCGCCACCTGGCCACCCAAACCTTGAACACATTTA 274

QY 417 CCCCGAGAGGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAGACCCAGCTGAAG 476

Db 275 CCCCGAGAGGCTGCTGCAGCAGATGAAGAGCTCTGACCGAGAACCCAGCTGAAG 334

QY 477 AAGCCATGAAGCTAAATAATCAAGCCATGAAGGAGATTGAGGAGCTTTCGGCTTGA 536

Db 335 AAGCCATGAAGCTAAATAATCAAGCCATGAAGGAGATTGAGGAGCTTTCGGCTTGA 394

QY 537 CAGAGAAAACAGAGGAAGAACCGCCAGTTTTTTTGTAGATACAGAGCAAGAGCAAGAGC 596

Db 395 CAGAGAAAACAGAGGAAGAACCGCCAGTTTTTTTGTAGATACAGAGCAAGAGCAAGAGC 454

QY 597 GTCTAATGGCTTTGAGTCATGAGATGAGAAATTGAGGAGAGCTTGGAAACTTAAAG 656

Db 455 GTCTAATGGCTTTGAGTCATGAGATGAGAAATTGAGGAGAGCTTGGAAACTTAAAG 514

QY 657 GGAATCAGAAAGGTCTCTGAGGAGCCCACTGTAGCTCCAGGCTTCCAGGGCCGAAG 716

Db	515	GGAAATCAGAAAGGCTCACTGAGGACCCCACTGATGACTCCAGGCTTCCAGGCGCCGAAG	574
Qy	717	CGGAGCAGGAAAGGACACAGCTCAGGAGCCCGAGGCTGAGGCTCAAGACGAGAGGCGAG	776
Db	575	CGGAGCAGGAAAGGACACAGCTCAGGAGCCCGAGGCTGAGGCTCAAGACGAGAGGCGAG	634
Qy	777	ACCTGTTGGGCAATCGTGCTGAACTGACGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAG	836
Db	635	ACCTGTTGGGCAATCGTGCTGAACTGACGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAG	694
Qy	837	ATTTCCTTTGTTCAAAATTAGGATGGCTGAAGGAGGAGCAGAGGCTCAGTAAAGAAATCA	896
Db	695	ATTTCCTTTGTTGAAATTAGGATGGCTGAAGGAGGAGCAGAGGCTCAGTAAAGAAATCA	754
Qy	897	AGCATAGTCTCGGGCCACGAGAACAGTCTCCACTGGCAGCGGCATTGCTAAAATATAGGA	956
Db	755	AGCATAGTCTCGGGCCACGAGAACAGTCTCCACTGGCAGCGGCATTGCTAAAATATAGGA	796
Qy	957	GCAGATCTCGAGATGGGGCCAAAGAATTAATTCGAACATGAGAGGTTAACTGTGAGCGAGC	1016
Db	797	GCAGATCTCGAGATGGGGCCAAAGAATTAATTCGAACATGAGAGGTTAACTGTGAGCGAGC	856
Qy	1017	TCTGCTGTGCTTAAGGAGGAGGATCAGAGGTGGAGAGACTTGAAGTTG	1067
Db	857	TCTGCTGTGCTTAAGGAGGAGGATCAGAGGTGGAGAGACTTGAAGTTG	907
RESULT 32			
DV778070		1201 bp	mRNA
LOCUS		Hw Fat 56_050802_F06 Bos taurus	CF-24-HW fat cDNA library Bos
DEFINITION		taurus cDNA, mRNA sequence.	
ACCESSION	DV778070.1	GI:82630946	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 833; Conservative			
Qy	1095	CAGATTTTGAAAGAAAAA	CAAGTAATCGTTCTGAGATTGAAACCCACGACAGAGGGGAGCA
Db	29	CTGATTTTGAAGAAAGCAAGGATCATTTCTGAGACTGACACCCCAACAGAGGAGCA	88

Matches		705;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1372	TCGTTTATCTAACA	AAAAAGCTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAAT	1431						
Db	1	TGTTTATCTAACA	AAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAAT	60						
QY	1432	GGACAGGCTAAAC	CAGAGGATGAAAGTCCAAATTAACCTGTGCTACAGATGACACAA	1491						
Db	61	GGACAGGCTAAAC	CAGAGGATGAAAGTCCAAATTAACCTGTGCTACAGATGACACAA	120						
QY	1492	CAAGCTCTTCAAG	AACAATAATATGCAATTTGAAACCAATTTGAGGAACCTAACCAAGAAAGA	1551						
Db	121	CAAGCTCTTCAAG	AACAATAATATGCAATTTGAAACCAATTTGAGGAACCTAACCAAGAAAGA	180						
QY	1552	GTCCAGAAAAGTG	GACGGCAGTCTGAAGGACTGAGGAACTGAGTGAAGAACTGGAACCTGGCAGA	1611						
Db	181	GTCCAGAAAAGTG	GACGGCAGTCTGAAGGACTGAGTGAAGAACTGGAACCTGGCAGA	240						
QY	1612	GAAGGCTCTGGCT	TCCAAACAGCTGCAAAATGGATGAAATGAAAGCAAAACCATTTGCCAAGCA	1671						
Db	241	GAAGGCTCTGGCT	TCCAAACAGCTGCAAAATGGATGAAATGAAAGCAAAACCATTTGCCAAGCA	300						
QY	1672	GGAAAGGACCTGG	AAACCATGACCATCTCAGGGCTCAGATGGAGTTTACTGTCTGA	1731						
Db	301	GGAAAGGACCTGG	AAACCATGACCATCTCAGGGCTCAGATGGAGTTTACTGTCTGA	360						
QY	1732	TTTTTCATGCTGA	AGAGCAGCAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCAAT	1791						
Db	361	TTTTTCATGCTGA	AGAGCAGCAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCAAT	420						
QY	1792	GCAGCTGGCAGTT	CTGCTGAAAGAGAAATGATGCTTTTTCGAAGACGGAGCAGGCTCCTT	1851						
Db	421	GCAGCTGGCAGTT	CTGCTGAAAGAGAAATGATGCTTTTTCGAAGACGGAGCAGGCTCCTT	480						
QY	1852	GATGGAGATCGAG	TCATGCTGGGCGAGAACAGATGACTCTGACGAGCAGGCTTACCT	1911						
Db	481	GATGGAGATCGAG	TCATGCTGGGCGAGAACAGATGACTCTGACGAGCAGGCTTACCT	540						
QY	1912	TGTTCAAAGAGAG	CTGAGCAGGAGCTGGCGGCAACAGCGGAATATTCGGATTCATTC	1971						
Db	541	TGTTCAAAGAGAG	CTGAGCAGGAGCTGGCGGCAACAGCGGAATATTCGGATTCATTC	600						
QY	1972	CTGCCCCAAGTGT	GAGAGGTTCTGCTGACATAGACACGTTTACAGATTCACGTGATGA	2031						
Db	601	CTGCCCCAAGTGT	GAGAGGTTCTGCTGACATAGACACGTTTACAGATTCACGTGATGA	660						
QY	2032	TTGCATCATTTAA	GTGTTGATGATACCTCCCAAAACCTGTTGGT	2077						
Db	661	TTGCATCATTTAA	GTGTTGATGATACCTCCCAAAACCTGTTGGT	706						

RESULT 35
BQ218575
LOCUS
DEFINITION BQ218575 878 bp mRNA linear EST 02-MAY-2002
AGENCY AGENT 7575749 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6047221
5', mRNA sequence.
ACCESSION BQ218575
VERSION BQ218575.1 GI:20399975
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 878)
NIH-MGC <http://mgc.nhl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13294 row: e column: 14
High quality sequence stop: 564.
Location/Qualifiers
1. 878
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6047221"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match		33.6%;	Score 697.4;	DB 3;	Length 878;
Best Local Similarity		97.8%;	Pred. No. 4.6e-169;		
Matches		749;	Conservative	0;	Mismatches 12; Indels 5; Gaps 4;
QY	297	CAGGAACTTCTGCA	ATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAAGGAGGACGAC	356	
Db	105	CAGGAACTTCTGCA	ATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAAGGAGGACGAC	164	
QY	357	CCAGTGAAGCACA	GAGAAATGGACCCCTCCCTGCGCCACCCAAACCTGGACACGTTTA	416	
Db	165	CCAGTGAAGCACA	GAGAAATGGACCCCTCCCTGCGCCACCCAAACCTGGACACGTTTA	224	
QY	417	CCCCGGAGGCTGT	CGCAGCAGATGAAAGAGCTCTGACCAAGAACCAACGAGCTGAAAG	476	
Db	225	CCCCGGAGGCTGT	CGCAGCAGATGAAAGAGCTCTGACCAAGAACCAACGAGCTGAAAG	284	
QY	477	AAGCCATGAAGCT	TAATAATCAAGCATGAAAGGGAGATTTGAGGAGCTTTCCGGCTCGA	536	
Db	285	AAGCCATGAAGCT	TAATAATCAAGCATGAAAGGGAGATTTGAGGAGCTTTCCGGCTCGA	344	
QY	537	CAGAAAAACAGAA	GAAGAAACGCGCCAGTTTTTTTGTAGATACAGAGCAAAAGCAAAAGAGC	596	
Db	345	CAGAAAAACAGAA	GAAGAAACGCGCCAGTTTTTTTGTAGATACAGAGCAAAAGCAAAAGAGC	404	
QY	597	GTCTAATGGCTTG	AGTCAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAA	656	
Db	405	GTCTAATGGCTTG	AGTCAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAA	464	
QY	657	GGAAATCAGAAAG	GTCACTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAAG	716	
Db	465	GGAAATCAGAAAG	GTCACTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAAG	524	
QY	717	CGGACAGGAAAGA	CCAGCTCAGGACCCAGGTTGAGGCTTCAAGCAGAGAAAGGAG	776	
Db	525	CGGACAGGAAAGA	CCAGCTCAGGACCCAGGTTGAGGCTTCAAGCAGAGAAAGGAG	584	
QY	777	ACCTGTTGGGCAT	CGTGTGAACTGCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG	836	
Db	585	ACCTGTTGGGCAT	CGTGTGAACTGCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG	644	
QY	837	ATTCTCTTTGTTT	GAATTAGGATGGCTGAAGGAGAAAGCAGAGGGTCAGTAAAAAGAAATCA	896	
Db	645	ATTCTCTTTGTTT	GAATTAGGATGGCTGAAGGAGAAAGCAGAGGGTCAGTAAAAAGAAATCA	704	
QY	897	AGCATAGTCTCTG	GGCCCGCAGGAAACAGTCTCCACTGGGACGGCATTTCTAAATATAGGA	956	
Db	705	AGCATAGTCTCTG	GGCCCGCAGGAAACAGTCTCCACTGGGACGGCATTTCTAAATATAGGA	764	
QY	957	GCAGATCTGCAG	ATGGGGCCAAG-AATTACTTTCGAACATG-AGGAGTTAACTGTGAGCCA	1014	
Db	765	GCAGATCTGCAG	ATGGGGCCAAG-AATTACTTTCGAACATG-AGGAGTTAACTGTGAGCCA	824	

QY	1015	G-CTCCGTGTCCTCAA--GGGAAGGGAATCAGAAGTGGAGAGA	1057
Db	825	GCCTCGTGCTGGGCCCTAAAGGGAAGGGGAATCCGAAGGGGGGAGA	870
RESULT	36		
CN309548			
LOCUS		696 bp mRNA linear EST 16-MAY-2004	
DEFINITION		L700582991172 GRN_PRENU Homo sapiens cDNA 5', mRNA sequence.	
ACCESSION		CN309548	
VERSION		1 GI:47325962	
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		1 (bases 1 to 696)	
AUTHORS		Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Quegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.	
TITLE		Transcriptome Characterization elucidates signaling networks that control human ES cell growth and differentiation	
JOURNAL		Nat. Biotechnol. 22 (6), 707-716 (2004)	
PUBMED		15146197	
COMMENT		Contact: Brandenberger R Regenerative Medicine Genon Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@genon.com Insert Length: 696 Std Error: 0.00. Location/Qualifiers	
FEATURES		source	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/tissue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"	
		/clone_lib="GRN PRENU"	
		/note="oligo dt primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic acid and mitogens."	
ORIGIN			
Query Match		33.4%; Score 694; DB 8; Length 696;	
Best Local Similarity		100.0%; Pred. No. 3.3e-168;	
Matches		694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	903	GTCTCGGGCCCACGAGAACAGTCTCCACTGGCACGGCAATTGCTTAATAATATAGGACAGAT	962
Db	1	GTCTCGGGCCCACGAGAACAGTCTCCACTGGCACGGCAATTGCTTAATAATATAGGACAGAT	60
QY	963	CTGCAGATGGGGCCAGAAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCTGC	1022
Db	61	CTGCAGATGGGGCCAGAAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCTGC	120
QY	1023	TGTGCTTAAGGGAGGGAATCAGAGGTGGAGAGACTTTGAATGGCTCACTCAAGGAGCCA	1082
Db	121	TGTGCTTAAGGGAGGGAATCAGAGGTGGAGAGACTTTGAATGGCTCACTCAAGGAGCCA	180
QY	1083	AACAAGAGTTTCAGATTTTGAAGAAAAGAAAAAACAAGTAATCGTTCGAGATTGAAACCCAGA	1142
Db	181	AAGAAAGAGTTTCAGATTTTGAAGAAAAGAAAAAACAAGTAATCGTTCGAGATTGAAACCCAGA	240
QY	1143	CAGAGGGGAGCACAGAGAAGAGATGATGAAGAGAAAGGCCCGGAGACTGTTGGAGCG	1202
Db	241	CAGAGGGGAGCACAGAGAAGAGATGATGAAGAGAAAGGCCCGGAGACTGTTGGAGCG	300
QY	1203	AGTGAAGACATGAACCTCCAGGTGA CATCTCTGTTTTAAGGAGCTTCAAGAGGCTCAT	1262

stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTGGGTG. The cell line was provided by Dr James Martin from University of Iowa. TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix TAG_LIB=UI-H-FL0 TAG_SEQ=GAGGTCGGTG"

ORIGIN

Query Match 33.4%; Score 693; DB 3; Length 733;
Best Local Similarity 100.0%; Pred. No. 6e-168;
Matches 693; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1385 AAAAGTTAGAGCTCAAGTGAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAA 1444
Db 733 AAAAAGTTAGAGCTCAAGTGAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAA 674
QY 1445 ACAGAGGATGAAGTCCAAATTAAGTGTCTACAGATGACACAAAGCTTCTTCAA 1504
Db 673 ACAGAGGATGAAGTCCAAATTAAGTGTCTACAGATGACACAAAGCTTCTTCAA 614
QY 1505 GAACATAATAATGCCATTGAAACAAATGAGGAACCTAAAGAAAGAGTCAGAAAAAGTG 1564
Db 613 GAACATAATAATGCCATTGAAACAAATGAGGAACCTAAAGAAAGAGTCAGAAAAAGTG 554
QY 1565 GACAGGCGAGTGTGAAGAACTGAGTGAAGAACTGGAACTGGCGAGAGAGGCTCTGGCT 1624
Db 553 GACAGGCGAGTGTGAAGAACTGAGTGAAGAACTGGAACTGGCGAGAGAGGCTCTGGCT 494
QY 1625 TCCAAACAGCTGCAATGATGAATGAGCAACCAATGCGCAACGAGAGAGGACCTG 1684
Db 493 TCCAAACAGCTGCAATGATGAATGAGCAACCAATGCGCAACGAGAGAGGACCTG 434
QY 1685 GAAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAA 1744
Db 433 GAAACCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAA 374
QY 1745 AGAGACGCGAGAGAGAAATTCATGAGGAAAGAGAGCAACTGGCAATTCGAGCTGGCAGTT 1804
Db 373 AGAGACGCGAGAGAGAAATTCATGAGGAAAGAGAGCAACTGGCAATTCGAGCTGGCAGTT 314
QY 1805 CTGCTGAAGAGAAATGATCTTTCGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1864
Db 313 CTGCTGAAGAGAAATGATCTTTCGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 254
QY 1865 AGTCGTGATGGGCGAGAAACAAGTGACTCTGACGAGAGGCTTACCTTGTTCAGAGAGGA 1924
Db 253 AGTCGTGATGGGCGAGAAACAAGTGACTCTGACGAGAGGCTTACCTTGTTCAGAGAGGA 194
QY 1925 GCTGAGGACAGGAGCTGGCGGAAACAGCGGAATATTCGAGTTCATTCCTGCCCCAAGTGT 1984
Db 193 GCTGAGGACAGGAGCTGGCGGAAACAGCGGAATATTCGAGTTCATTCCTGCCCCAAGTGT 134
QY 1985 GGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCAGTGTGATGGATTCATTTAA 2044
Db 133 GGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCAGTGTGATGGATTCATTTAA 74
QY 2045 GTGTTGATGTATCACCTCCCCAAAACCTGTGTGT 2077
Db 73 GTGTTGATGTATCACCTCCCCAAAACCTGTGTGT 41

RESULT 38

BQ879423
LOCUS BQ879423
DEFINITION AGENCOURT 8119646 Lupski dorsal root ganglion Homo sapiens cDNA
clone IMAGE:6177842 5', mRNA sequence.
ACCESSION BQ879423
VERSION BQ879423.1 GI:22271431
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 904)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13556 row: h column: 03
High quality sequence stop: 636.
FEATURES
Location/Qualifiers
1..904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177842"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGGCTCCG-3' and
5'-GACTAGTTCTAGATCGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match 33.3%; Score 690.8; DB 3; Length 904;
Best Local Similarity 99.3%; Pred. No. 2.4e-167;
Matches 714; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 297 CAGGAACCTTCGCAATGTCCTCAATCAACCTCTCAGCTGCTCACTGAAAAGGAGGACGCC 356
Db 184 CAGGAACCTTCGCAATGTCCTCAATCAACCTCTCAGCTGCTCACTGAAAAGGAGGACGCC 243
QY 357 CCAGTGAAGCACAGGAATGGACCCCGCCACCTGGCCACCCAACTGGACAGCTTTA 416
Db 244 CCAGTGAAGCACAGGAATGGACCCCGCCACCTGGCCACCCAACTGGACAGCTTTA 303
QY 417 CCCCAGGAGCTGCTGCAGCAGATGAAAGAGCTCTGACCAAGAACCAACCTGGACAGTAA 476
Db 304 CCCCAGGAGCTGCTGCAGCAGATGAAAGAGCTCTGACCGAGAACCAACCTGGACAGTAA 363
QY 477 AAGCCATGAAGCTAAATTAATCAAGCCATGAAGGAGATTTGAGGAGCTTTCCGCTGGA 536
Db 364 AAGCCATGAAGCTAAATTAATCAAGCCATGAAGGAGATTTGAGGAGCTTTCCGCTGGA 423
QY 537 CAGGAAACAGAGGAAGAACCCAGTTTGTGATACAGAGCAAGAAAGCAAGAGAGC 596
Db 424 CAGGAAACAGAGGAAGAACCCAGTTTGTGATACAGAGCAAGAAAGCAAGAGAGC 483
QY 597 GTCTAATGGCTTGAAGTCAATGAGAAATGAGAAATGAGAAAGCTTGGAAAACCTAAAAG 656
Db 484 GTCTAATGGCTTGAAGTCAATGAGAAATGAGAAATGAGAAAGCTTGGAAAACCTAAAAG 543
QY 657 GGAAATCAGAAAGGTCATCTGAGGACCCCACTGATGATCCAGGCTTCCAGGGCCGAG 716


```

UI-H-FEO-bby-d-03-0-UI 3', mRNA sequence.
ACCESSION BQ773529
VERSION BQ773529.1 GI:21982005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 723)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
    source
    1..723
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="UI-H-FEO-bby-d-03-0-UI"
    /tissue_type="Chondrosarcoma Cell line"
    /dev_stage="Adult"
    /lab_host="DH10B (Life Technologies)"
    /clone_lib="NCI-CGAP_FEO"
    /note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoR I; Site 2: Not I; NCI-CGAP_FEO is
a cDNA library containing the following tissue(s): a pool
of 3 chondrosarcoma cell lines (grade 2) The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCTACGGAC. The cell lines was provided by Dr James Martin
of University of Iowa.
TAG_TISSUE=human grade 2 chondrosarcoma cell line pool
TAG_LIB=UI-H-FEO
TAG_SEQ=CGCTACGGAC"

ORIGIN
Query Match 32.8%; Score 682; DB 3; Length 723;
Best Local Similarity 99.9%; Pred. No. 4.3e-165;
Matches 682; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1395 ACCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAACAGAGGATG 1454
Db |-----|
723 AGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAACAGAGGATG 664
QY 1455 AAAAGTCCAAATTAAGTGTCTACAGATGACACACACAGCTTCTTCAGACACATAATA 1514
Db |-----|
663 AAAAGTCCAAATTAAGTGTCTACAGATGACACACACAGCTTCTTCAGACACATAATA 604
QY 1515 ATGCAATTGAAAACAAATTGAGGAACCTAACAGAAAAGAGTCAGAAAAGTGGACAGGCAG 1574
Db |-----|
603 ATGCAATTGAAAACAAATTGAGGAACCTAACAGAAAAGAGTCAGAAAAGTGGACAGGCAG 544
QY 1575 TGCTGAAGGAACGTAGTGAAGAACTGGAACTGGCAGAGAGAGGCTCTGGCTTCCAAACAGC 1634
Db |-----|
543 TGCTGAAGGAACGTAGTGAAGAACTGGAACTGGCAGAGAGAGGCTCTGGCTTCCAAACAGC 484

```

```

QY 1635 TGCAAAATGGATGAATGAAGCAAAACCAATTGCCAAGCAGAAAGAGACCTGGAAACCATGA 1694
Db |-----|
483 TGCAAAATGGATGAATGAAGCAAAACCAATTGCCAAGCAGAAAGAGACCTGGAAACCATGA 424
QY 1695 CCATCCTCAGGGCTCAGATGGAAGTTTACTGTTCTGTATTTTTCATGCTGAAAGAGCAGCA 1754
Db |-----|
423 CCATCCTCAGGGCTCAGATGGAAGTTTACTGTTCTGTATTTTTCATGCTGAAAGAGCAGCA 364
QY 1755 GAGGAAAAATTCATGAGGAAAAAGGCAACATGGCAATTCGACCTGCGACGCTCTGCTGAAG 1814
Db |-----|
363 GAGGAAAAATTCATGAGGAAAAAGGCAACATGGCAATTCGACCTGCGACGCTCTGCTGAAG 304
QY 1815 AGAATGATGCTTTTGAAGACGAGGACGAGCTCCTTGATGGAGATGCAGAGTCGTCATG 1874
Db |-----|
303 AGAATGATGCTTTTGAAGACGAGGACGAGCTCCTTGATGGAGATGCAGAGTCGTCATG 244
QY 1875 GGGCGAGAACCAAGTGACTCTGACCAAGAGGCTTACCTTGTTCAAAGAGAGAGTGAGGACA 1934
Db |-----|
243 GGGCGAGAACCAAGTGACTCTGACCAAGAGGCTTACCTTGTTCAAAGAGAGAGTGAGGACA 184
QY 1935 GGGACTGCGGCAACAGGCGGAATATTCGATTTCATCTGCCCCCAAGTGTGGAGAGGTTTC 1994
Db |-----|
183 GGGACTGCGGCAACAGGCGGAATATTCGATTTCATCTGCCCCCAAGTGTGGAGAGGTTTC 124
QY 1995 TGCTTGACATAGACACAGTTACAGATTACAGTTCAGTTCATCATTTTAAGTTGTGATGT 2054
Db |-----|
123 TGCTTGACATAGACACAGTTACAGATTACAGTTCAGTTCAGTTCATCATTTTAAGTTGTGATGT 64
QY 2055 ATCACTCTCCCAAAACTGTGGT 2077
Db |-----|
63 ATCACTCTCCCAAAACTGTGGT 41

```

```

RESULT 41
BQ776232/c
LOCUS BQ776232
DEFINITION UI-H-FHO-bcn-i-23-0-UI-s1 NCI CGAP FHO Homo sapiens cDNA clone
UI-H-FHO-bcn-i-23-0-UI 3', mRNA sequence.
ACCESSION BQ776232
VERSION BQ776232.1 GI:21984708
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 726)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

```

```

FEATURES
    source
    1..726
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="UI-H-FHO-bcn-i-23-0-UI"
    /tissue_type="Human Chondrosarcoma Cell Line"
    /dev_stage="Adult"
    /lab_host="DH10B (Life Technologies)"
    /clone_lib="NCI-CGAP_FHO"
    /note="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a

```


modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP FHO is a cDNA library containing the following
tissue(s): Human Grade 1 Chondrosarcoma Cell line The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is AGAATCCGCG. The cell line was provided by Dr
James Martin from University of Iowa
TAG TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
Chondrosarcoma
TAG LIB=UI-H-FHO
TAG_SEQ=AGAATCCGCG"

ORIGIN

Query Match 32.8%; Score 680.8; DB 3; Length 726;
Best Local Similarity 99.6%; Pred. No. 8.7e-165;
Matches 682; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1393 AGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAACAGAGGA 1452
DB 726 AGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAACAGAGGA 667

QY 1453 TGAAGATGCCAAATTAACCTGCTACAGATGACACACAAAGCTTCTTCAAGAACATAA 1512
DB 666 TGAAGATGCCAAATTAACCTGCTACAGATGACACACAAAGCTTCTTCAAGAACATAA 607

QY 1513 TAATGCAATTGAAACCAATTTAGGAACATAACAGAAAGAGTTCAGAAAGAGTGACAGGGC 1572
DB 606 TAATGCAATTGAAACCAATTTAGGAACATAACAGAAAGAGTTCAGAAAGAGTGACAGGGC 547

QY 1573 AGTCTGGAAGAACTAGTGAAAAAATCGGAACTGGCAGAGAGAGGCTCTGCTCCAAACA 1632
DB 546 AGTCTGGAAGAACTAGTGAAAAAATCGGAACTGGCAGAGAGAGGCTCTGCTCCAAACA 487

QY 1633 GCTGCAAAATGGAATGAAGCAACCACTGTCAGAGCAGGAGAGAGGCTGGAACACCT 1692
DB 486 GCTGCAAAATGGAATGAAGCAACCACTGTCAGAGCAGGAGAGAGGCTGGAACACCT 427

QY 1693 GACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCAGC 1752
DB 426 GACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCAGC 367

QY 1753 GAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCAATTCAGCTGGCAGTTCTGCTGAA 1812
DB 366 GAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCAATTCAGCTGGCAGTTCTGCTGAA 307

QY 1813 AGAGATGATGCTTTGAGAGACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1872
DB 306 AGAGATGATGCTTTGAGAGACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 247

QY 1873 TGGGGGCGAGAAAGTGAATGACTCTGACAGCAGGAGGCTTACCTTGTTCAGAGAGGAGCTCAGGA 1932
DB 246 TGGGGGCGAGAAAGTGAATGACTCTGACAGCAGGAGGCTTACCTTGTTCAGAGAGGAGCTCAGGA 187

QY 1933 CAGGAGCTGGCGGCAACAGCGGAATATTCGATTCATTCCTGCCCCAAGTGTGGAGAGGT 1992
DB 186 CAGGAGCTGGCGGCAACAGCGGAATATTCGATTCATTCCTGCCCCAAGTGTGGAGAGGT 127

QY 1993 TCTGCTGATAGACACGTTTACAGATTCACGTTGATGGATTGATTCATTTAAGTGTGAT 2052
DB 126 TCTGCTGATAGACACGTTTACAGATTCACGTTGATGGATTGATTCATTTAAGTGTGAT 67

QY 2053 GTATCACTCTCCCAAACTGTTGGT 2077
DB 66 GTATCACTCTCCCAAACTGTTGGT 42

RESULT 42
BU628287/c

LOCUS
DEFINITION

742 bp mRNA linear EST 23-SEP-2002
UI-H-FGO-bde-1-07-0-UI.s1 NCI CGAP EN1_2 Homo sapiens cDNA clone
UI-H-FGO-bde-1-07-0-UI.3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU628287

UI-H-FGO-bde-1-07-0-UI.s1

UI-H-FGO-bde-1-07-0-UI.3'

BU628287

BU628287.1

GI:23294501

EST

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 742)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: James Martin

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1..742

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-FGO-bde-1-07-0-UI"

/tissue_type="Enchondroma cell line"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP EN1_2"

/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

NCI CGAP EN1_2 is a cDNA library containing the following

tissue(s): Enchondroma cell line (2 cell lines). The

library was constructed according to Bonaldo, Lennon and

Soares, Genome Research, 6:791-806, 1996. First strand

cDNA synthesis was primed with an oligo-dT primer

containing a Not I site. Double stranded cDNA was ligated

to an EcoR I adaptor, digested with Not I, and cloned

directionally into pT7T3-Pac vector. The oligonucleotide

used to prime the synthesis of first-strand cDNA contains

a library tag sequence that is located between the Not I

site and the (dT)18 tail. The sequence tag for this

library is CCGTCACTC. The cell lines was provided by Dr

James Martin from University of Iowa.

TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)

TAG_LIB=UI-H-FGO

TAG_SEQ=CGTCACTC"

ORIGIN

Query Match 32.8%; Score 680.4; DB 3; Length 742;
Best Local Similarity 99.7%; Pred. No. 1.1e-164;
Matches 692; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1377 ATACTAACAAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC 1436
DB 742 ATACTAACAAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC 683

QY 1437 AGGCTAAAAACAGAGGATGAAAAGTCCAAATTAACCTGTGTACAGATGACACACAAAGC 1496
DB 682 AGGCTAAAAACAGAGGATGAAAAGTCCAAATTAACCTGTGTACAGATGACACACAAAGC 623

QY 1497 TTCTTCAAGAACATATAATGCAATTGAAAAACAATTGAGGAACCTAACAGAAAAGATGTCAG 1556
DB 622 TTCTTCAAGAACATATAATGCAATTGAAAAACAATTGAGGAACCTAACAGAAAAGATGTCAG 563

```
QY 1557 AAAAAAGTGACAGGCGAGTGCTGAAGGAACCTGAGTGAAAAAAGCTGGAACCTGGCAGAGAAGG 1616
|
|
|
Db 562 AAAAAAGTGACAGGCGAGTGCTGAAGGAACCTGAGTGAAAAAAGCTGGAACCTGGCAGAGAAGG 503
|
|
|
QY 1617 CTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCTATGCCAAGCAGGAAG 1676
|
|
|
Db 502 CTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCTATGCCAAGCAGGAAG 443
|
|
|
QY 1677 AGGACCTGGAACCAATGACCATCTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTC 1736
|
|
|
Db 442 AGGACCTGGAACCAATGACCATCTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTC 383
|
|
|
QY 1737 ATGCTGAAAGACGACGAGAGAGAAATTCATGAGAAAGAGCAACTGGCATTGCGAGC 1796
|
|
|
Db 382 ATGCTGAAAGACGACGAGAGAGAAATTCATGAGAAAGAGCAACTGGCATTGCGAGC 323
|
|
|
QY 1797 TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTCGAAGACGGAGCGAGGAGTCTTTGATGG 1856
|
|
|
Db 322 TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTCGAAGACGGAGCGAGGAGTCTTTGATGG 263
|
|
|
QY 1857 AGATGACAGTGTCTATGGGGGAGAAACGAGTGACTCTGACACGAGGCTTACCTTGTTTC 1916
|
|
|
Db 262 AGATGACAGTGTCTATGGGGGAGAAACGAGTGACTCTGACACGAGGCTTACCTTGTTTC 203
|
|
|
QY 1917 AAGAGGAGCTGAGACAGGAGTGGCGGCAACAGCGGAATATTCGATTTCATCTCGCC 1976
|
|
|
Db 202 AAGAGGAGCTGAGGACAGGAGTGGCGGCAACAGCGGAATATTCGATTTCATCTCGCC 143
|
|
|
QY 1977 CC-AAAGTGTGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCACTGATGGATTGC 2035
|
|
|
Db 142 CC-AAAGTGTGAGAGGTTCTGCTGACATAGACAGCTTACAGATTCACTGATGGATTGC 83
|
|
|
QY 2036 ATCATTAAAGTGTGATGATATCACCTCCCAAAA 2069
|
|
|
Db 82 ATCATTAAAGTGTGATGATATCACCTCCCAAAA 49
|
|
|
```

```
RESULT 43
CD245471
LOCUS
DEFINITION
AGENCOURT 14098672 NIH_MGC_181 Homo sapiens cDNA clone
IMAGE:30376546 5', mRNA sequence.
CD245471
CD245471.1 GI:31005935
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 871)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabps@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM439 row: f column: 11
High quality sequence stop: 670.
Location/Qualifiers
1..871
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30376546"
/tissue_type="White Matter"
```

```
/dev stage="Unknown"
/lab host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone lib="NIH_MGC_181"
/ncbi vector="PCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
(note=Vector: PCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match 32.7%; Score 679.8; DB 5; Length 871;
Best Local Similarity 99.4%; Pred. No. 1.7e-164;
Matches 693; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 297 CAGGAACCTTCGCAATGTCCTCAACCTCTCAGCTGCTCACTGAAAGAGGAGCAGCC 356
|
|
|
Db 124 CAGGAACCTTCGCAATGTCCTCAACCTCTCAGCTGCTCACTGAAAGAGGAGCAGCC 183
|
|
|
QY 357 CCAGTGAAGACACAGGAAATGGACCCCTCCACCTGGCCCCACCCAAACCTGGACACGTTTA 416
|
|
|
Db 184 CCAGTGAAGACACAGGAAATGGACCCCTCCACCTGGCCCCACCCAAACCTGGACACGTTTA 243
|
|
|
QY 417 CCCGGAGAGCTGCTGACAGATGAAGAGCTCTCTGACCAAGAACCAACAGCTGAAAG 476
|
|
|
Db 244 CCCGGAGAGCTGCTGACAGATGAAGAGCTCTCTGACCAAGAACCAACAGCTGAAAG 303
|
|
|
QY 477 AAGCCATCAAGCTAAATCAAGCCATGAAGGGAGATTTGAGGAGCTTTTCGGCTTGA 536
|
|
|
Db 304 AAGCCATCAAGCTAAATCAAGCCATGAAGGGAGATTTGAGGAGCTTTTCGGCTTGA 363
|
|
|
QY 537 CAGAAAAACAGAAAGAAAGCCACCTGTTTGTAGATACAGAGCAAAAGAAAGAGC 596
|
|
|
Db 364 CAGAAAAACAGAAAGAAAGCCACCTGTTTGTAGATACAGAGCAAAAGAAAGAGC 423
|
|
|
QY 597 GTCTAATGGCTTGTAGTATGAGAAATGAAGAGAGCTTTGAGAAACTTAAAG 656
|
|
|
Db 424 GTCTAATGGCTTGTAGTATGAGAAATGAAGAGAGCTTTGAGAAACTTAAAG 483
|
|
|
QY 657 GGAATCAGAAAGGTCATCTGAGGACCCACCTGATGACTCCAGGCTTCCAGGGCCGAAG 716
|
|
|
Db 484 GGAATCAGAAAGGTCATCTGAGGACCCACCTGATGACTCCAGGCTTCCAGGGCCGAAG 543
|
|
|
QY 717 CGGACAGGAAAAAGGACCCAGCTCAGGACCCAGGTGGTGAGGCTACAAGCAGAGAGGCGAG 776
|
|
|
Db 544 CGGACAGGAAAAAGGACCCAGCTCAGGACCCAGGTGGTGAGGCTACAAGCAGAGAGGCGAG 603
|
|
|
QY 777 ACCTGTTGGGCAATCGTGTCTGAACTGACAGCTCAAGCTGAACTCCAGCGGCTCTCAGAG 836
|
|
|
Db 604 ACCTGTTGGGCAATCGTGTCTGAACTGACAGCTCAAGCTGAACTCCAGCGGCTCTCAGAG 663
|
|
|
QY 837 ATTCCTTTGTTCAAAATTAGGATGGCTGAAGAGAGGAGAGGGTCAAGTAAAGAAATCA 896
|
|
|
Db 664 ATTCCTTTGTTGAAATTAGGATGGCTGAAGAGAGGAGAGGGTCAAGTAAAGAAATCA 723
|
|
|
QY 897 AGCATAGTCTCGGGCCACGAGAAACAGTCTCCACTGGCACCGGCTTGTCTAAATATAGA 956
|
|
|
Db 724 AGCATAGTCTCGGGCCACGAGAAACAGTCTCCACTGGCACCGGCTTGTCTAAATATAGA 783
|
|
|
QY 957 GCAGATCTGCAGATGGGG--CCAAGAAATTACTTCGAA 991
|
|
|
Db 784 GCAGATCTGCAGATGGGGCCCGAGGAATTAATCTTCGAA 820
|
|
|
```

```
RESULT 44
BG031732
LOCUS
DEFINITION
602300889F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4402102 5',
mRNA sequence.
ACCESSION
BG031732
VERSION
BG031732.1 GI:12422313
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
```

```
EST 24-JAN-2001
linear
clone
cDNA
clone
IMAGE:4402102 5',
mRNA
sequence.
Homo sapiens
Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 740)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE NIH-MGC http://mgc.nci.nih.gov/
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov

Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L14M10110 row: b column: 23
High quality sequence stop: 675.

FEATURES

source location/Qualifiers
1..740
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4402102"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_87"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 32.3%; Score 670.8; DB 2; Length 740;
Best Local Similarity 99.6%; Pred. No. 3.4e-162;
Matches 683; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1097 GATTTTGAAGAAACAACTGCTCTGAGATTGAAACCCAGACAGAGGGGAGCACA 1156
DB 10 GCTTTTGAAGAAACAACTGCTCTGAGATTGAAACCCAGACAGAGGGGAGCACA 69
QY 1157 GAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTGGAAGCGAGTGGAGCACTG 1216
DB 70 GAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTGGAAGCGAGTGGAGCACTG 129
QY 1217 AACTCCAGTGACATCTCTGTTTAAGGACCTTCAAGAGCTCATCAAAACTCAGCGAA 1276
DB 130 AACTCCAGTGACATCTCTGTTTAAGGAGCTTCAAGAGCTCATCAAAACTCAGCGAA 189
QY 1277 GCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTGAGGCCCTTGAAGGAAATTTCT 1336
DB 190 GCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTGAGGCCCTTGAAGGAAATTTCT 249
QY 1337 GCAATTCATCAGAGTTGAATGAAACCAAGAGCTGTTTATATCAACAAAAAGTTAGAG 1396
DB 250 GCAATTCATCAGAGTTGAATGAAACCAAGAGCTGTTTATATCAACAAAAAGTTAGAG 309
QY 1397 CTACAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAGGATGAA 1456
DB 310 CTACAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAGGATGAA 369
QY 1457 AAGTCCAAATTAAGTGTGTACAGATGACACAAAGCTTCTTCAAGAAACATAATAT 1516
DB 370 AAGTCCAAATTAAGTGTGTGTACAGATGACACAAAGCTTCTTCAAGAAACATAATAT 429
QY 1517 GCATTGAAAACAATTGAGGAAGTAAACAGAAAAAGAGTCAGAAAAAGTGGACAGGGCAGTG 1576
DB 430 GCATTGAAAACAATTGAGGAAGTAAACAGAAAAAGAGTCAGAAAAAGTGGACAGGGCAGTG 489
QY 1577 CTGAAGAACTGAGTGAAGAACTGGAACTGGGAGAGAGGCTCTGGCTTCCAAACAGCTG 1636
DB 490 CTGAAGAACTGAGTGAAGAACTGGAACTGGGAGAGAGGCTCTGGCTTCCAAACAGCTG 549

QY 1637 CAAATGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAGGACCTGGAAACCATGACC 1696
DB 550 CAAATGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAGGACCTGGAAACCATGACC 609
QY 1697 ATCCTCAGGCTCAGATGGAAGTTTACTGTTTCTGATTTTCATGCTGAAAGCAGCAGAG 1756
DB 610 ATCCTCAGGCTCAGATGGAAGTTTACTGTTTCTGATTTTCATGCTGAAAGCAGCAGAG 668
QY 1757 GAGAAATTCATGAGGAAAGGAGCA 1782
DB 669 GAGAAATTCATGAGGAAAGGAGCA 694

RESULT 45

AU122473
LOCUS AU122473
DEFINITION AU122473 MAMMAL Homo sapiens cDNA clone MAMMAL1002441 5', mRNA
sequence.

ACCESSION AU122473
VERSION AU122473.1 GI:10937743
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 895)
AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Sugano, S., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.

TITLE Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
JOURNAL Genome Res. 16 (1), 55-65 (2006)
PUBMED 16344560
COMMENT Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES
source location/Qualifiers
1..895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMAL1002441"
/tissue_type="mammary gland"
/clone_lib="MAMMAL"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 32.0%; Score 664.2; DB 1; Length 895;
Best Local Similarity 98.9%; Pred. No. 1.9e-160;
Matches 699; Conservative 0; Mismatches 5; Indels 3; Gaps 3;
QY 297 CAGGAACCTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAAGGAGCAGCC 356
DB 131 CAGGAACCTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAAGGAGCAGCC 190
QY 357 CCAGTGAAGACAGAGAAATGGACCCGCCACCTGGCCACCCAAACCTTGGACACGTTTA 416
DB 191 CCAGTGAAGACAGAGAAATGGACCCGCCACCTGGCCACCCAAACCTTGGACACGTTTA 250

417 CCCCAGGAGGCTGCTGCAGCAGATCAAGAGCTCTGACCAAGAACACACAGCTGAAAG 476
|||||
Db CCCCAGGAGGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAACACACAGCTGAAAG 310
|||||
QY AAGCCATGAAGCTAAATATCAAGCCATGAAGAGGAGATTTTCAGAGGCTTCCGGCTCGA 536
Db AAGCCATGAAGCTAAATATCAAGCCATGAAGAGGAGATTTTCAGAGGCTTCCGGCTCGA 370
|||||
QY CAGAGAAACAGAGGAGAACCCAGTCTTTTTCAGATACAGAGCAAGAGCAAGAGAGCC 596
Db CAGAGAAACAGAGGAGAACCCAGTCTTTTTCAGATACAGAGCAAGAGCAAGAGAGCC 430
|||||
QY GTCTAATGCGCTTGAGTCAATGAGATGAGAAATGAAGGAGAGCTTGAAAACTTAAAG 656
Db GTCTAATGCGCTTGAGTCAATGAGATGAGAAATGAAGGAGAGCTTGAAAACTTAAAG 490
|||||
QY GGAATACAGAAAGGTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAG 716
Db GGAATACAGAAAGGTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAG 550
|||||
QY CGGACAGAGAAAGGACAGCTCAGAGCCAGGCTGAGGCTACAGAGAGAGAGGCGAG 776
Db CGGACAGAGAAAGGACAGCTCAGAGCCAGGCTGAGGCTACAGAGAGAGAGGCGAG 610
|||||
QY ACTGTTGGGCTCGTGTCTGAACTGCAGCTCAAGCTCAAGCTCCAGCGGCTCCTCAGAG 836
Db ACTGTTGGGCTCGTGTCTGAACTGCAGCTCAAGCTCAAGCTCCAGCGGCTCCTCAGAG 670
|||||
QY ATTCTTTGTTGAAATAGAGTGGCTGAAGAGAGAGAGAGAGAGGCTCAGTAAAGAAATCA 896
Db ATTCTTTGTTGAAATAGAGTGGCTGAAGAGAGAGAGAGAGGCTCAGTAAAGAAATCA 730
|||||
QY AGCATAGTCTCGGGCCACGAGACAGTCTCCACTGGCAGCGATGCTTAAATATAGA 956
Db AGCATAGTCTCGGGCCACGAGACAGTCTCCACTGGCAGCGATGCTTAAATATAGA 789
|||||
QY GCAGATCTCAGATGGGGCCAGAAATTTACTTCGAAATCAGGAGGTTA 1003
Db NGATCTCAGATGGGGCCAGAAATTTACTTCGAAATCAGGAGGTTA 834

RESULT 46
BI062019/c
LOCUS IL3-UT0117-130301-499-E06 UT0117 Homo sapiens cDNA, mRNA linear EST 15-JUN-2001
DEFINITION
ACCESSION BI062019
VERSION BI062019.1 GI:14469546
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 664)
AUTHORS Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,D.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL3&t2=IL3-UT0117-
130301-499-E06&t3=2001-03-13&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 663.
Location/Qualifiers
FEATURES
source
1..664
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UT0117"
/note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 31.8%; Score 659.8; DB 2; Length 664;
Best Local Similarity 99.5%; Pred. No. 2.4e-159;
Matches 661; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1205 GTGGAAGCACTGAACCTCCAGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACA 1264
Db GTGGAAGCACTGAACCTCCAGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACA 605
QY 1265 AAATCAGCGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTGAGGCCCTTGA 1324
Db AAATCAGCGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTGAGGCCCTTGA 545
QY 1325 AGGAAAATTTCTGCATTTCCATCAGAGTTGAATGAAGAGAGAGCTTGTATATCTAAC 1384
Db AGGAAAATTTCTGCATTTCCATCAGAGTTGAATGAAGAGAGAGCTTGTATATCTAAC 485
QY 1385 AAAAAGTTAGAGCTACAAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAAGCTAAA 1444
Db AAAAAGTTAGAGCTACAAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAAGCTAAA 425
QY 1445 ACAGAGGATGAAAAGTCCAAATTAATCTGTCTACAGATGACACACAAGAGCTTCTTCAA 1504
Db ACAGAGGATGAAAAGTCCAAATTAATCTGTCTACAGATGACACACAAGAGCTTCTTCAA 365
QY 1505 GAACATATATGCAATTCGAAACATTCAGGAACTTAAACAAGAAAAGAGTTCAGAAAAGTG 1564
Db GAACATATATGCAATTCGAAACATTCAGGAACTTAAACAAGAAAAGAGTTCAGAAAAGTG 305
QY 1565 GACAGGCGAGTGTCTGAAGAACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCTGGCT 1624
Db GACAGGCGAGTGTCTGAAGAACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCTGGCT 245
QY 1625 TCCAAACAGCTGCAAAATGAGATGAAGCAAAACATTCGCCAAGCAGGAGAGACCTTG 1684
Db TCCAAACAGCTGCAAAATGAGATGAAGCAAAACATTCGCCAAGCAGGAGAGACCTTG 185
QY 1685 GAAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAA 1744
Db GAAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAA 125
QY 1745 AGAGCAGCGAGAGAGAAAATTCATGAGAAAAGAGCAACTGGCATTGCGAGTGGCAGTT 1804
Db AGAGCAGCGAGAGAGAAAATTCATGAGAAAAGAGCAACTGGCATTGCGAGTGGCAGTT 65
QY 1805 CTGCTGAAAGAGATGATGCTTTCAGACGAGGAGGAGCACTCTTGTATGAGATGCGAG 1864
Db CTGCTGAAAGAGATGATGCTTTCAGACGAGGAGGAGCACTCTTGTATGAGATGCGAG 5
QY 1865 AGTC 1868
Db 4 AGTC 1

```

RESULT 47
DN110950
LOCUS      111027 MARC 4PIG Sus scrofa cDNA 5', mRNA linear EST 14-FEB-2005
DEFINITION
ACCESSION  DN110950
VERSION     DN110950.1 GI:59789935
KEYWORDS   EST.
SOURCE      Sus scrofa (pig)
ORGANISM    Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.

REFERENCE
1 (bases 1 to 891)
AUTHORS    Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
            Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE      Porcine EST collection using a normalized library constructed from
            embryos representing early developmental stages
JOURNAL    Unpublished (2003)
COMMENT    Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: HHY8005 row: H column: 8
            Seq primer: GTAATACGACTCACTATAGGG.
            Location/Qualifiers
                . . 891
                /organism="Sus scrofa"
                /mol_type="mRNA"
                /db_xref="taxon:9823"
                /tissue_type="pooled"
                /lab_host="DH108"
                /clone_lib="MARC 4PIG"
                /notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
                Library made with combined RNA from day-10, day-13,
                day-15, day-25, and day-30 whole embryos."

FEATURES
source
    Query Match      31.8%; Score 659.8; DB 9; Length 891;
    Best Local Similarity 84.0%; Pred. No. 2.6e-159;
    Matches 764; Conservative 0; Mismatches 127; Indels 18; Gaps 1;

QY 867 GAGAACGAGAGGTCAGTAAAGAAATCAAGCAATAGCATAGTCTGGGCCACGAGAACAGTCT 926
      |||||
Db 1 GAGAACGAGATGAGCAATGAGGAAATCAAGCAAGTCTGGGCCCATAGAACTGATT 60

QY 927 CCACTGCACGGCATTGTCTAATATAGGACGAGATCTGCAGATGGGCCCAAGATTACT 986
      |||||
Db 61 CCATTGACAC-----GAGCAATCTGCAGAAAGTACCAGGAATTATT 102

QY 987 TCGAATGAGGAGTAACTGTGAGCAGCTCTCTGTGCTTAAGGAGGAAATCAGA 1046
      |||||
Db 103 TGGAAATTTGAGGAATTAATCTGTGAGCCAGCTCTCTGTGCTTAAGGAGGAAATCAGA 162

QY 1047 AGGTGAGAGACTTGGAATTTGCACTCAAGAGAGCCCAAGAAAGAGTTTCAGATTTGAAA 1106
      |||||
Db 163 AGGTGAGAGACTTGAATTCGCCCTCAAGGAGCCCAAGAAAGAAATTTAGATTTGAAA 222

QY 1107 AGAAACAAGTATCTCTGAGATTGAACCCAGACAGAGGGGAGCAGACAGAAAGAGA 1166
      |||||
Db 223 AGAAAGCAAGGATCTGTTCTGAGACTGAGACCCAGACAGAAAGAGCAAAAGAAACAGAGA 282

QY 1167 ATGATGAAGAGAAAGCCCGAGACTGTTGGAAGCGAAGTGGAGCACTGAACCTCCAGG 1226
      |||||
Db 283 AAGAGAGGAGAAAGCCCAAGAACTGTTGGAAGTGAAGTGGAAATTTAAACCTTCAGG 342

QY 1227 TGCATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGCTAA 1286
      |||||

```

```

Db 343 TGACRAACCCCTGTTTAAAGAGCTTTCAGGAGGCTCACAGAAACTCAGTGAAGCTGAGTCA 402
QY 1287 TGAAGAAGAGAGACTTCAAGAAAAAGTGTGAGCCCTTGAAGAGGAAAAATTTCTGCAATTCAT 1346
      |||||
Db 403 TGAAGAAGAGAGACTTCAAGAAAAATGTGAGGCACTTGAAGAGGAAAAATTTCTGCAACCCAT 462
      |||||
QY 1347 CAGAGTTGAATGAAAAAGCAAGAGCTTGTATTACTAACAAAAAGTTAGAGCTACAAGTGG 1406
      |||||
Db 463 CAGAGCTGAATGAAAAAGCAAGAGCTTCTTTATATAACAAAAAGTTGGAGCTCCAAGTGG 522
      |||||
QY 1407 AAGCATGCTATCAGAAATCAAAATGGAACAGCTTAAACAGAGGATGAAGAACTCCAAAT 1466
      |||||
Db 523 AAGCATGAGATCAGAAATTTAAATGGAGCAAGCCAAACAGAAAGGAAAAAGTCCAAAT 582
      |||||
QY 1467 TAACTGTGTACAGATGACACACAAAGCTTCTTCAAGAACATAATATATGCAATGAAAA 1526
      |||||
Db 583 TAACTACTTCAAGTTGACCCACACAGGCTTCTTCAAGNATACATATATGCACTGAAAA 642
      |||||
QY 1527 CAATTGAGAACTTAACAAGAAAGAGTCAAGAAAAAGTGAAGGAGGAGTCTGCAAGGAAC 1586
      |||||
Db 643 CAATTGAGAACTTGAAGAAAGAGAGAGTCTGAAAAAGTGGATAGGTGGTGTCTGAGGAAC 702
      |||||
QY 1587 TGAGTGAATAACTGGAACTGGCAGAGAGAGGCTCTGCTTCCAAACAGCTGCAATGGATG 1646
      |||||
Db 703 TGAATGGAAGCTGGAAATGGCAGAGAGGCTTCCAGAGGCTTCCAAAGAGCTTCAAAATGGATG 762
      |||||
QY 1647 AAATGAAGCAAACTATTGCAAGCAGAAAGAGAGCTTGGAAACCATGACCATCTCAGGG 1706
      |||||
Db 763 AGATGAAGCAGACCATTTGCAAGCAGAGAGAGGACTTGGAAACCATGCTGTTCTCAGGG 822
      |||||
QY 1707 CTCAGATGGAAGTTTACTGTTCTGATTTTTCATGCTGAAAGAGCAGGAGAGAGAAATTC 1766
      |||||
Db 823 CTCAGATGAGGTATATCTGTTCTGACTTTTCATGCTGAAAGAGCAGCAGCAAGAGAGATTC 882
      |||||
QY 1767 ATGAGGAAA 1775
      |||||
Db 883 ATGAAGAAA 891

RESULT 48
DN111300/c
LOCUS      1110411 MARC 4PIG Sus scrofa cDNA 3', mRNA linear EST 14-FEB-2005
DEFINITION
ACCESSION  DN111300
VERSION     DN111300.1 GI:59789966
KEYWORDS   EST.
SOURCE      Sus scrofa (pig)
ORGANISM    Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.

REFERENCE
1 (bases 1 to 875)
AUTHORS    Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
            Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE      Porcine EST collection using a normalized library constructed from
            embryos representing early developmental stages
JOURNAL    Unpublished (2003)
COMMENT    Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: HHY8005 row: H column: 8
            Seq primer: TAGAAGCAGCAGTCCAGG.
            Location/Qualifiers
                . . 875
                /organism="Sus scrofa"
                /mol_type="mRNA"
                /db_xref="taxon:9823"

```

```
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4P1G"
/Note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN

Query Match          31.2%; Score 649; DB 9; Length 875;
Best Local Similarity 87.1%; Pred. No. 1.6e-156;
Matches 738; Conservative 0; Mismatches 100; Indels 9; Gaps 2;

QY 1240 TAAGAGCTTCAAGAGGCTCATCAAACTCAGCGAAGCTGAGCTAAATGAAGAAGAGACT 1299
      |||||
Db 875 TAAGAGCTTCAAGAGGCTCACAGAACTCAGTGAAGCTGAGCTCATGAAGAAGAGACT 816
      |||||

QY 1300 TCAAGAAAGTGTAGGCCCTTGAAGAGAAATTCGCAATTCATCAGAGTTGAATGA 1359
      |||||
Db 815 TCAAGAAATGTCAAGCACTTGAAGAGAAATTCGCAACCCCATCAGAGCTGAATGA 756
      |||||

QY 1360 AAGCAAGAGCTTGTATTACTAAACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATC 1419
      |||||
Db 755 AAGCAAGAGCTTCTTTATATACAAAAGTTGAGCTCCAAGTGGAAAGCATGAGATC 696
      |||||

QY 1420 AGAAATCAAAATGGACAGGCTTAACACAGAGGATGAAAAGTCCAAATTAACCTGCTACA 1479
      |||||
Db 695 AGAAATTAATGGAGCAAGCCAAACACAGAGGAAAGTCCAAATTAACCTACTACA 636
      |||||

QY 1480 GATGACACACAGAGCTTCTTCAAGAACATATATGATTTGAAAACATTTGAGAACT 1539
      |||||
Db 635 GTTGACCCACACAGGCTTCTTCAAGAAATCAATATGCACTGAAAACATTTGAGAACT 576
      |||||

QY 1540 AACAGAAAGAGTGCAGAAAAGTGCACAGGCGAGTCTGGAAGGAAGTCAAGTGAAGAACT 1599
      |||||
Db 575 GAAAGAGAGAGTCTGAAAAGTGGATAAGTGTGCTGCAAGAACTGAATGGAAGCT 516
      |||||

QY 1600 GGAAGTGGCAGAGAGGCTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAC 1659
      |||||
Db 515 GGAAGTGGCAGAGAGGCTCTGGCTTCCAAAGCAGCTCCAAATGGATGAGATGAAGCAGAC 456
      |||||

QY 1660 CATTCGCAAGCAGGAAGAGGACCTGGAACCATGACCATCTCCAGGGCTCAGATGGAAGT 1719
      |||||
Db 455 CATTCGCAAGCAGGAAGAGGACCTGGAACCATGCTGCTCTCAGGGCTCAGATGGAAGT 396
      |||||

QY 1720 TTACTGTTCTGATTTTCATGCTGAAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGGA 1779
      |||||
Db 395 ATACTGTTCTGACTTTTCATGCTGAAAGAGCAGCGAGAGAGAAATTCATGAGGAAAGGA 336
      |||||

QY 1780 GCAACTGGCAATTCAGCTGGCAGTTCTGCTGAAAGAGAAATGATGCTTTCCAGAGCGGAG- 1838
      |||||
Db 335 GCAACTGGCAATTCAGCTGGCAGTTTCTGCTGAAAGAGCAATGCTTTTGAAGAGGGAGC 276
      |||||

QY 1839 --GCAGGAGTCTTCAATGAGATGAGAGTGTGATGCGGGCGAGAACAGAGTACTCTGA 1896
      |||||
Db 275 CAGCAGGCAATCTTGTGGAATGCAAGCCGCTCATGGGGCAAGCAAGTGTGCTGA 216
      |||||

QY 1897 CAGCAGGCTTACTTGTTCAGAGAGGAGTCAAGGAGGAGTGGCGGCAACAG----- 1951
      |||||
Db 215 CAGCAGGCTTGTCTTGTTCAGAGGAGGAGTGAAGATAGAACTGGCTCAGCAGCAACA 156
      |||||

QY 1952 -CGGAATATTCGATTCATTCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATGACAC 2010
      |||||
Db 155 ACAGAAATATTCAAATTCATTCTTGCCCCCAATGTGGAGAAAGTTCTGCTGACATGATAC 96
      |||||

QY 2011 GTTACAGATTCAGGATGATGATCATTTTAAGTGTGATGATATCACTCCCCCAAC 2070
      |||||
Db 95 ACTACTGATTCAGGTTACGGAGTGCATCTTTTAAGTGTGCTGACGTTTTTACTCCCCCAAC 36
      |||||

QY 2071 TGTGGT 2077
      |||||
Db 35 TGTGGT 29
      |||||
```

RESULT 49
BU626293/c
LOCUS
DEFINITION
UI-H-DF0-bet-h-21-0-UI s1 NCI CGAP_DF0 Homo sapiens cDNA clone
UI-H-DF0-bet-h-21-0-UI 3', mRNA sequence.
ACCESSION
BU626293
VERSION
BU626293.1 GI:23292508
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 693)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-bet-h-21-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_DF0"
/note="Organ: Bone; Vector: pTVT3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP_DF0 is a cDNA library containing the following
tissue(s): Subchondral Bone. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pTVT3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GTTAAGCGTC.
TAG TISSUE=Subchondral bone
TAG_LIB=UI-H-DF0
TAG_SEQ=GTTAAGCGTC"

Query Match 31.0%; Score 643; DB 3; Length 693;
Best Local Similarity 99.8%; Pred. No. 5.4e-155;
Matches 654; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1423 AATCAAAATGGAACAGGCTTAAACAGAGATGAAAAGTCCAAATTAACCTGTGCTACAGAT 1482
 |||||
Db 693 AATCAAAATGGAACAGGCTTAAACAGAGATGAAAAGTCCAAATTAACCTGTGCTACAGAT 634
 |||||
QY 1493 GACACACACAGCTCTTTCAGAACATAATATCATTTGAAAACAAATTCAGGAACTAAC 1542
 |||||
Db 633 GACACACACAGAG- TTCTTCAAGAACATAATATGCAATTTGAAAACAAATTCAGGAACTAAC 575
 |||||
QY 1543 AAGAAAAGAGTCAGAAAAAGTGGACAGGCGCAGTGTGGAAGGAACCTGAGTGAAAACTGGA 1602
 |||||
Db 574 AAGAAAAGAGTCAGAAAAAGTGGACAGGCGCAGTGTGGAAGGAACCTGAGTGAAAACTGGA 515
 |||||
QY 1603 ACTGCGAGAGAGGCTCTGGCTTCCAAACAGCTGCAATGGATGAAATGAAGCAACCAAT 1662

Db 514 ACTGGCAGAGAGGCTGGCTTCCAAACAGCTGCAAAATGATGAATGAAGCAAAACCAT 455
QY TGCCCAAGCAGGAGAGGACCTGGAAACCATGACCATCTCAGGCTCAGATGGAAGTTTA 1722
Db 454 TGCNAGCAGGAGAGGACCTGGAAACCATGACCATCTCAGGCTCAGATGGAAGTTTA 395
QY CTGTTCTGATTTTTCATGCTGAAAGAGCAGGAGAGAGAAATTCATGAGGAAAGAGCA 1782
Db 394 CTGTTCTGATTTTTCATGCTGAAAGAGCAGGAGAGAGAAATTCATGAGGAAAGAGCA 335
QY 1783 ACTGGCATTGCGAGCTGCTGCTGAAAGAGATGCTTCCAAACAGGAGGCGAG 1842
Db 334 ACTGGCATTGCGAGCTGCTGCTGAAAGAGATGCTTCCAAACAGGAGGCGAG 275
QY 1843 GCAGTCTGATGAGATGAGAGTCTGCTGAGGAGGAGCAACAGTGCATCTGACCAAGCA 1902
Db 274 GCAGTCTGATGAGATGAGAGTCTGCTGAGGAGGAGCAACAGTGCATCTGACCAAGCA 215
QY 1903 GGCTTACCTGTTTCAAGAGGAGCTGAGGACAGGAGTGGCGGCAACAGCGGAATATTC 1962
Db 214 GGCTTACCTGTTTCAAGAGGAGCTGAGGACAGGAGTGGCGGCAACAGCGGAATATTC 155
QY 1963 GATTATCTGCTGCCCAAGTGTGGAGAGGTTCTGCTGACATACAGACGTTACAGATTCA 2022
Db 154 GATTATCTGCTGCCCAAGTGTGGAGAGGTTCTGCTGACATACAGACGTTACAGATTCA 95
QY 2023 CGTGATGGATTGCATCATTTAAGTGTGATGATGATACCTCCCAAAACTGTTGT 2077
Db 94 CGTGATGGATTGCATCATTTAAGTGTGATGATGATACCTCCCAAAACTGTTGT 40

RESULT 50
BQ893002
LOCUS AGENCOURT 8118169 lupski_dorsal_root_ganglion Homo sapiens cDNA
DEFINITION clone IMAGE:6181131 5', mRNA sequence.
ACCESSION BQ893002
VERSION BQ893002.1 GI:22285016
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 905)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM13565 Row: a Column: 04
High quality sequence stop: 619.
Location/Qualifiers
1..905
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6181131"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski dorsal root ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.

FEATURES
Source

Directionally cloned using the following adaptors:
5'-TCCAGCCAGCGTCCG-3' and
5'-GACTAGTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN
Query Match 30.8%; Score 640.4; DB 3; Length 905;
Best Local Similarity 98.2%; Pred. No. 2.8e-154;
Matches 647; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 297 CAGGAACTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAGAGGACAGCC 356
Db 143 CAGGAACTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAGAGGACAGCC 202
QY 357 CAGTGAAGACACAGGAAATGGAACCCCTGCGCCCAACCCAACTCTGGACACGTTTA 416
Db 203 CCAGTGAAGACACAGGAAATGGAACCCCTGCGCCCAACCCAACTCTGGACACGTTTA 262
QY 417 CCCGGAGGAGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAACCCAGCTGAAG 476
Db 263 CCCGGAGGAGCTGCTGCAGCAGATGAAGAGCTCTGACCAAGAACCCAGCTGAAG 322
QY 477 AAGCCATGAAGCTTAAATTAATCAAGCCATGAAGAGGAGATTTGAGGAGCTTTTGGCCCTGGA 536
Db 323 AAGCCATGAAGCTTAAATTAATCAAGCCATGAAGAGGAGATTTGAGGAGCTTTTGGCCCTGGA 382
QY 537 CAGAGAAACAGAGAGGAGAAACGCCAGTTTTTTTTCAGATACAGAGCAAGAAAGCAAAAGAGC 596
Db 383 CAGAGAAACAGAGAGGAGAAACGCCAGTTTTTTTTCAGATACAGAGCAAGAAAGCAAAAGAGC 442
QY 597 GTCTAATGCGCTTGAGTCAATGAGAAATGAGAAATTTGAAGAGAGCTTTGAAAACATAAAG 656
Db 443 GTCTAATGCGCTTGAGTCAATGAGAAATTTGAAGAGAGCTTTGAAAACATAAAG 502
QY 657 GGAATCAGAAAGGTTCATCTGAGGAGCCCACTGATGATCCAGGCTTCCCGAGGCCGGAAG 716
Db 503 GGAATCAGAAAGGTTCATCTGAGGAGCCCACTGATGATCCAGGCTTCCCGAGGCCGGAAG 562
QY 717 CGGAGCAGGAAAGGAGCCAGCTCAGGAGCCCACTGAGGCTGAGGCTTACAGCAGAGAGGCGAG 776
Db 563 CGGAGCAGGAAAGGAGCCAGCTCAGGAGCCCACTGAGGCTTACAGCAGAGAGGCGAG 622
QY 777 ACCTGTTGGGCATCGTGTCTGAACTCAGCTCAAGCTGAACCTCAGCGGCTCTCTAGAAG 836
Db 623 ACCTGTTGGGCATCGTGTCTGAACTCAGCTCAAGCTGAACCTCAGCGGCTCTCTAGAAG 682
QY 837 ATTCTCTTCTGTAATTTAGGATGGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA 896
Db 683 ATTCTCTTCTGTAATTTAGGATGGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA 742
QY 897 AGCATAGTCTCTGGGCCACGAGAAAGTCTCCACTGGCAGCGCATTTCTCTAAATATAGG 955
Db 743 AGCATAGTCTCTGGGCCACGAGAAAGTCTCTCCACTGGCAGCGCATTTCTCAGATGGG 801

Search completed: May 30, 2006, 05:35:07
Job time : 8911 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2006, 09:22:57 ; Search time 664 Seconds
(without alignments)
5852.851 Million cell updates/sec

Title: SEQ1-458A
Perfect score: 2077
Sequence: 1 atcccggtcgaggattctt.....acctcccaaaactgttgt 2077

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues
Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Issued Patents NA:*

- 1: /EMC Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
- 2: /EMC Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
- 3: /EMC Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
- 4: /EMC Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
- 5: /EMC Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
- 6: /EMC Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
- 7: /EMC Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
- 8: /EMC Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
- 9: /EMC Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
- 10: /EMC Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES	
Result No.	Score	Query Match	Length DB ID	Description	
1	665.2	32.0	710 4	US-09-297-648-4781	Sequence 4781, Ap
2	424.6	20.4	444 3	US-09-621-976-802	Sequence 802, App
3	134.8	6.5	428 3	US-09-621-976-10957	Sequence 10957, A
C 4	80.8	3.9	7218 2	US-08-232-463-14	Sequence 14, Appl
	77.8	3.7	1874 3	US-09-253-701-2	Sequence 2, Appli
6	76.6	3.7	1994 5	US-10-156-306B-8014	Sequence 8014, Ap
7	76.6	3.7	1994 5	US-09-543-679A-2716	Sequence 2716, Ap
8	76.6	3.7	2093 3	US-09-646-403-1	Sequence 1, Appli
9	76.6	3.7	2034 3	US-09-646-403-2	Sequence 2, Appli
10	76.6	3.7	2035 3	US-09-863-049B-3	Sequence 3, Appli
11	76.6	3.7	8631 5	US-09-543-679A-2717	Sequence 2717, Ap
12	49.8	2.4	1827 3	US-09-270-767-1308	Sequence 1308, Ap
13	49.8	2.4	1827 3	US-09-270-767-16590	Sequence 16590, A
14	46.8	2.3	5433 3	US-08-529-329-1	Sequence 1, Appli
15	46.4	2.2	614 3	US-09-902-540-1318	Sequence 1318, Ap
16	46.2	2.2	2719 3	US-09-949-016-3860	Sequence 3860, Ap
17	46	2.2	627 3	US-09-710-279-693	Sequence 693, App
18	46	2.2	627 3	US-09-710-279-1325	Sequence 1325, Ap
19	46	2.2	633 3	US-09-134-001C-578	Sequence 578, App
C 20	46	2.2	2939 3	US-09-710-279-4039	Sequence 4039, Ap
	46	2.2	3691 3	US-09-710-279-3338	Sequence 3338, Ap
C 21	46	2.2	705 3	US-09-270-767-5061	Sequence 5061, Ap
	45.2	2.2	705 3	US-09-270-767-20343	Sequence 20343, A
22	45.2	2.2	705 3	US-09-270-767-20343	Sequence 20343, A
23	45.2	2.2	705 3	US-09-270-767-20343	Sequence 20343, A

24	44.6	2.1	696	3	US-09-461-697-193	Sequence 193, App
25	44.6	2.1	699	3	US-09-461-697-191	Sequence 191, App
26	44.6	2.1	717	3	US-09-461-697-189	Sequence 189, App
27	44.6	2.1	774	3	US-09-461-697-187	Sequence 187, App
28	44.6	2.1	819	3	US-09-461-697-185	Sequence 185, App
29	44.6	2.1	1669	3	US-09-461-697-184	Sequence 184, App
30	44.6	2.1	3113	2	US-08-993-228-20	Sequence 20, Appl
31	44.4	2.1	612	3	US-09-902-540-1357	Sequence 1357, Ap
32	44.4	2.1	1893	3	US-09-248-796A-6172	Sequence 6172, Ap
33	44	2.1	851	3	US-09-270-767-8286	Sequence 8286, Ap
C 34	44	2.1	851	3	US-09-270-767-23568	Sequence 23568, A
35	44	2.1	23106	3	US-09-863-049B-1	Sequence 1, Appli
36	43.8	2.1	12695	3	US-09-949-016-16775	Sequence 16775, A
37	43.6	2.1	601	3	US-09-949-016-77355	Sequence 77355, A
C 38	43.6	2.1	55841	3	US-09-949-016-16602	Sequence 16602, A
39	43.4	2.1	1149	3	US-09-336-115C-23	Sequence 23, Appl
40	43.4	2.1	3305	3	US-09-710-279-4002	Sequence 4002, Ap
C 41	43	2.1	1141	3	US-09-806-708B-22	Sequence 22, Appl
42	43	2.1	2330	2	US-08-557-128-3	Sequence 3, Appli
43	43	2.1	2330	3	US-09-242-690A-34	Sequence 34, Appl
44	43	2.1	2330	3	US-09-908-855-34	Sequence 34, Appl
45	42.8	2.1	1685	3	US-10-104-047-1957	Sequence 1957, Ap
46	42.6	2.1	345	3	US-09-614-221A-426	Sequence 426, App
47	42.6	2.1	1039	3	US-09-502-540-1280	Sequence 1280, Ap
48	42.4	2.0	154626	3	US-09-949-016-14000	Sequence 14000, A
C 49	41.8	2.0	396	3	US-09-640-173-53	Sequence 53, Appl
C 50	41.8	2.0	396	3	US-09-713-550-53	Sequence 53, Appl
C 51	41.8	2.0	396	3	US-09-825-294-53	Sequence 53, Appl
C 52	41.8	2.0	396	3	US-09-370-966-53	Sequence 53, Appl
C 53	41.8	2.0	601	3	US-09-949-016-173109	Sequence 173109, A
54	41.8	2.0	20229	3	US-09-949-016-16649	Sequence 16649, A
55	41.6	2.0	601	3	US-09-949-016-146976	Sequence 146976, A
56	41.6	2.0	2947	3	US-09-949-016-815	Sequence 815, App
57	41.6	2.0	38611	3	US-09-949-002-649	Sequence 649, App
58	41.6	2.0	38611	3	US-09-949-002-809	Sequence 809, App
C 59	41.2	2.0	66085	3	US-09-949-016-13292	Sequence 13292, A
60	41	2.0	289	3	US-09-007-005-17	Sequence 17, Appl
61	41	2.0	289	3	US-09-244-796-17	Sequence 17, Appl
C 62	41	2.0	601	3	US-09-949-016-193621	Sequence 193621, A
C 63	41	2.0	601	3	US-09-949-016-193622	Sequence 193622, A
C 64	41	2.0	601	3	US-09-949-016-193623	Sequence 193623, A
C 65	41	2.0	601	3	US-09-949-016-193624	Sequence 193624, A
66	41	2.0	634	3	US-08-881-094-16	Sequence 16, Appl
67	41	2.0	268449	3	US-09-949-016-17244	Sequence 17244, A
68	40.8	2.0	478	3	US-09-621-976-10407	Sequence 10407, A
C 69	40.8	2.0	601	3	US-09-949-016-53576	Sequence 53576, A
70	40.8	2.0	2169	3	US-09-434-408-3	Sequence 3, Appli
71	40.8	2.0	3610	4	US-09-880-107-2406	Sequence 2406, Ap
72	40.8	2.0	5058	3	US-09-889-595-1	Sequence 1, Appli
73	40.8	2.0	5058	3	US-09-899-595-1	Sequence 10574, A
74	40.6	2.0	1122	3	US-09-270-767-10574	Sequence 2, Appli
75	40.6	2.0	5919	3	US-08-875-435B-2	Sequence 12707, A
C 76	40.4	1.9	190078	3	US-09-949-016-12707	Sequence 17026, A
C 77	40.4	1.9	190078	3	US-09-949-016-17026	Sequence 15868, A
C 78	40.4	1.9	374159	3	US-09-949-016-15868	Sequence 173108, A
C 79	40.2	1.9	601	3	US-09-949-016-173108	Sequence 845, App
80	40.2	1.9	2049	3	US-09-107-532A-845	Sequence 845, App
81	40.2	1.9	2273	3	US-09-949-016-3986	Sequence 3986, Ap
82	40.2	1.9	2523	3	US-09-861-451A-29	Sequence 29, Appl
C 83	40.2	1.9	90541	3	US-09-759-359A-3	Sequence 3, Appli
C 84	40.2	1.9	90541	3	US-10-207-973-3	Sequence 3, Appli
C 85	40.2	1.9	90541	5	US-10-799-676-3	Sequence 3, Appli
86	40.2	1.9	186734	3	US-09-949-016-14870	Sequence 14870, A
87	40.2	1.9	193689	3	US-09-949-016-12350	Sequence 12350, A
88	40.2	1.9	193689	3	US-09-949-016-13088	Sequence 13088, A
C 89	40	1.9	399	3	US-09-621-976-8976	Sequence 8976, Ap
C 90	40	1.9	601	3	US-09-949-016-43337	Sequence 43337, A
C 91	40	1.9	601	3	US-09-949-016-43338	Sequence 43338, A
C 92	40	1.9	601	3	US-09-949-016-43566	Sequence 43566, A
C 93	40	1.9	601	3	US-09-949-016-43567	Sequence 43567, A
C 94	40	1.9	601	3	US-09-949-016-43795	Sequence 43795, A
C 95	40	1.9	601	3	US-09-949-016-43796	Sequence 43796, A
96	40	1.9	1411	3	US-09-370-838-154	Sequence 154, App

```
97 40 1.9 1411 3 US-09-854-133-154 Sequence 154, App
98 40 1.9 2001 3 US-09-949-016-2801 Sequence 2801, Ap
99 40 1.9 8257 3 US-09-595-684B-30 Sequence 30, Appl
100 40 1.9 8503 3 US-09-620-312D-130 Sequence 130, App
101 40 1.9 113701 3 US-09-949-016-13214 Sequence 13214, A
102 39.8 1.9 601 3 US-09-949-002-1592 Sequence 1592, Ap
103 39.8 1.9 601 3 US-09-949-002-1593 Sequence 1593, Ap
104 39.8 1.9 601 3 US-09-949-002-10498 Sequence 10498, A
105 39.8 1.9 601 3 US-09-949-002-10499 Sequence 10499, A
106 39.8 1.9 3565 3 US-09-973-278-939 Sequence 929, App
107 39.8 1.9 38855 3 US-09-973-278-930 Sequence 930, App
108 39.8 1.9 40908 3 US-09-949-002-845 Sequence 845, App
109 39.8 1.9 40944 3 US-09-949-002-603 Sequence 603, App
110 39.6 1.9 3489 2 US-08-728-323A-1 Sequence 1, Appli
111 39.6 1.9 3489 3 US-09-298-568-1 Sequence 1, Appli
112 39.6 1.9 3489 3 US-09-410-393-1 Sequence 1, Appli
113 39.6 1.9 3489 3 US-09-894-273-1 Sequence 1, Appli
114 39.6 1.9 26050 3 US-09-949-016-17449 Sequence 17449, A
115 39.6 1.9 32207 3 US-08-770-379-20 Sequence 20, Appl
116 39.6 1.9 32207 3 US-08-757-669A-20 Sequence 20, Appl
117 39.6 1.9 32207 3 US-09-230-371A-20 Sequence 20, Appl
118 39.6 1.9 43192 3 US-09-949-016-15466 Sequence 15466, A
119 39.4 1.9 1469 3 US-09-949-018-5398 Sequence 5398, Ap
120 39.4 1.9 4875 2 US-08-460-739-1 Sequence 1, Appli
121 39.4 1.9 79595 3 US-09-949-016-15318 Sequence 15318, A
122 39.4 1.9 148794 3 US-09-949-016-12751 Sequence 12751, A
123 39.2 1.9 277 3 US-09-007-005-3 Sequence 3, Appli
124 39.2 1.9 277 3 US-09-244-796-3 Sequence 3, Appli
125 39 1.9 2277 2 US-08-676-967-2 Sequence 2, Appli
126 39 1.9 2277 2 US-08-676-974-2 Sequence 2, Appli
127 39 1.9 2277 2 US-09-098-487-2 Sequence 2, Appli
128 39 1.9 2771 3 US-10-104-047-154 Sequence 154, App
129 39 1.9 4167 3 US-09-973-278-700 Sequence 700, App
130 39 1.9 346112 3 US-09-949-016-13165 Sequence 13165, A
131 38.8 1.9 601 3 US-09-949-016-76961 Sequence 76961, A
132 38.8 1.9 601 3 US-09-949-016-76962 Sequence 76962, A
133 38.8 1.9 601 3 US-09-949-016-76963 Sequence 76963, A
134 38.8 1.9 601 3 US-09-949-016-76964 Sequence 76964, A
135 38.8 1.9 1266 3 US-09-248-796A-5218 Sequence 5218, Ap
136 38.8 1.9 1266 3 US-09-248-796A-3540 Sequence 3540, Ap
137 38.8 1.9 1960 2 US-08-533-306A-1 Sequence 1, Appli
138 38.8 1.9 1960 2 US-08-742-923A-1 Sequence 1, Appli
139 38.8 1.9 2353 3 US-10-104-047-1120 Sequence 1120, Ap
140 38.8 1.9 2680 2 US-08-533-306A-5 Sequence 5, Appli
141 38.8 1.9 2680 2 US-08-742-923A-5 Sequence 5, Appli
142 38.8 1.9 2887 2 US-08-533-306A-3 Sequence 3, Appli
143 38.8 1.9 2887 2 US-08-742-923A-3 Sequence 3, Appli
144 38.8 1.9 4206 3 US-09-750-590A-3 Sequence 3, Appli
145 38.8 1.9 4730 3 US-09-750-590A-1 Sequence 1, Appli
146 38.8 1.9 5574 3 US-09-917-254-40 Sequence 40, Appli
147 38.8 1.9 5937 4 US-09-927-597-1 Sequence 1, Appli
148 38.8 1.9 5937 4 US-09-927-597-1 Sequence 1, Appli
149 38.8 1.9 6861 3 US-09-949-016-1240 Sequence 1240, Ap
150 38.8 1.9 6861 3 US-09-949-016-1241 Sequence 1241, Ap
```

ALIGNMENTS

```
RESULT 1
US-09-297-648-4781
; Sequence 4781, Application US/09297648
; Patent No. 6964868
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominquez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
```

```
; APPLICANT: Pot, David
; APPLICANT: Kaskan, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297,648
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4781
; LENGTH: 710
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-297-648-4781
```

```
Query Match 32.0%; Score 665.2; DB 4; Length 710;
Best Local Similarity 98.7%; Pred No. 1.3e-176;
Matches 681; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1009 GAGCCAGCTCTCTGTCCTTAAGGGAAGGGGAATCAGAGGTGGAGAGACTTGAAGTTCG 1068
Db 22 GATCCCTCGATTCTGTCGCTTAAGGGAAGGGGAATCAGAGGTGGAGAGACTTGAAGTTCG 81

QY 1069 ACTCAGAGGCGCAAGAAAGTTTCAGATTTTCAAAAGAAAACAAGTAATCGTTCTGA 1128
Db 82 ACTCAGAGGCGCAAGAAAGTTTCAGATTTTCAAAAGAAAACAAGTAATCGTTCTGA 141

QY 1129 GATTGAAACCCACAGACAGAGGGGAGGACACAGAGAAAGAGATGATGAAGAGAAAGGCCCGGA 1188
Db 142 GATTGAAACCCACAGACAGAGGGGAGGACACAGAGAAAGAGATGATGAAGAGAAAGGCCCGGA 201

QY 1189 GACTGTGTGAAGCGAAGTGGAAAGCACTGAACCTCCAGGTGCACTCTCTTTTAAGAGCT 1248
Db 202 GACTGTGTGAAGCGAAGTGGAAAGCACTGAACCTCCAGGTGCACTCTCTTTTAAGAGCT 261

QY 1249 TCAAGAGGCTCATACAAAACCTCAGCGAGCTCAGCTTAATGAAGAGAGAGACTTCAAGAAA 1308
Db 262 TCAAGAGGCTCATACAAAACCTCAGCGAGCTCAGCTTAATGAAGAGAGAGACTTCAAGAAA 321

QY 1309 GTGTGAGGCCCTTGAAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAAGCAAGA 1368
Db 322 GTGTGAGGCCCTTGAAAGGAAAAATTCGCAATTCATCAGAGTTGAATGAAAGCAAGA 381

QY 1369 GCTTGTATTATACACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAA 1428
Db 382 GCTTGTATTATACACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAA 441

QY 1429 AATGACAGGCTAAACACAGAGGATGAAAGTCCCAATTAACCTGTGCTCAGATGACACA 1488
Db 442 AATGACAGGCTAAACACAGAGGATGAAAGTCCCAATTAACCTGTGCTCAGATGACACA 501
```

Qy 1489 CAACAAGCTTCTTCAAGAACATATAATGCAATGAAACAATGAGGAACATAACAAGAAA 1548
Db 502 CAACAAGCTTCTTCAAGAACATATAATGCAATGAAACAATGAGGAACATAACAAGAAA 561
Qy 1549 AGAGTCAGAAAAGTGGACAGGCGAGTCTGAAGAACTGAGTGAACAACTGGAACTGGC 1608
Db 562 AGAGTCAGAAAAGTGGACAGGCGAGTCTGAAGAACTGAGTGAACAACTGGAACTGGC 621
Qy 1609 AGAAGAGCTCTGGCTTCCAAACACAGCTGCAATGGAATGAAGCAAACTGTCGCAA 1668
Db 622 AGAAGAGCTCTGGCTTCCAAACACAGCTGCAATGGAATGAAGCAAACTGTCGCAA 681
Qy 1669 GCAGAGAGAGAGCTGGAACCATGACCAT 1698
Db 682 GCAGAGAGAGG-CCTGGAACCATGACCAT 710
RESULT 2
US-09-621-976-802
; Sequence 802, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 802
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 191..442
US-09-621-976-802

Query Match 20.4%; Score 424.6; DB 3; Length 444;
Best Local Similarity 98.0%; Pred. No. 4.7e-109;
Matches 438; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
Qy 1093 TTCAGATTTTCAAAAGAAAACAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAG 1152
Db 1 TTCAGATTTTCAAAAGAAAACAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAG 60
Qy 1153 CACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTGGAAGCGAAAGTGAAGC 1212
Db 61 CACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTGGAAGCGAAAGTGAAGC 120
Qy 1213 ACTGAACCTCCAGTGACATCTCTTTTAGGAGCTTCAAGAGCTCATACAAACTCAG 1272
Db 121 ACTGAACCTCCAGTGACATCTCTTTTAGGAGCTTCAAGAGCTCATACAAACTCAG 180
Qy 1273 CGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTGAGCCCTTGAAGGAAAA 1332
Db 181 CGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTGAGCCCTTGAAGGAAAA 237
Qy 1333 TTCTGCAATTCATCAGAGTTGAAATGAAAAGCAAGAGCTGTTTATATCTAACCAAAAGTT 1392
Db 238 TTCTGCAATTCATCAGAGTTGAAATGAAAAGCAAGAGCTGTTTATATCTAACCAAAAGTT 297
Qy 1393 AGAGCTCAGAGTGAAGAGCATGCTATCAGAAATCAAAATGAAACAGGCTAAACAGAGGA 1452
Db 298 AGAGCTCAGAGTGAAGAGCATGCTATCAGAAATCAAAATGAAACAGGCTAAACAGAGGA 357
Qy 1453 TGAAGCTCAAAATTAAGTCTGCTACAGATGACACAAAGCTTCTTCAAGACATAA 1512
Db 358 TGAAGCTCAAAATTAAGTCTGCTACAGATGACACAAAGCTTCTTCAAGACATAA 417

Qy 1513 TAATGCATTGAAAAACAATTTGAGGAAC 1539
Db 418 TAATGCATTGAAAAACAATTTGAGGAAC 444
RESULT 3
US-09-621-976-10957
; Sequence 10957, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 10957
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-10957
Query Match 6.5%; Score 134.8; DB 3; Length 428;
Best Local Similarity 98.6%; Pred. No. 1.5e-27;
Matches 136; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1708 TCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGTAAGAGACGAGAGAGAAAAATTC 1767
Db 235 TAAGATGGAAGTTTACTGTTCTGATTTTCATGCTGTAAGAGACGAGAGAGAAAAATTC 294
Qy 1768 TGAGGAAAAGAGCAACTGCGATTGCGAGTGGCGAGTTCTGCTGAAAGAGATGATGCTTT 1827
Db 295 TGAGGAAAAGAGCAACTGCGATTGCGAGTGGCGAGTTCTGCTGAAAGAGATGATGCTTT 354
Qy 1828 CGAAGACGAGGAGCGCA 1845
Db 355 CGAAGACGAGGAGCGGTA 372

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991

COMPOSITIONS, KIT & METHOD FOR TREATMENT
OF AIRWAY DISORDERS ASSOCIATED WITH
BRONCHOCONSTRICTION, LUNG INFLAMMATION,
NUMBER OF SEQUENCES: 3111
CORRESPONDENCE ADDRESS:
ADDRESSEE: EPIGENESIS PHARMACEUTICALS, INC.
STREET: 7 Clarke Drive
CITY: Cranbury
STATE: NJ
COUNTRY: USA
ZIP: 08512
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: N/A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/543,679A
FILING DATE: 13-Apr-2000
CLASSIFICATION: UNKNOWN
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/127,958
FILING DATE: 1998-08-03
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: EPI-0067191b
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-409-3035
TELEFAX: 413-254-9245
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2716:
SEQUENCE CHARACTERISTICS:
LENGTH: 1994 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2716

Query Match 3.7%; Score 76.6; DB 5; Length 1994;
Best Local Similarity 57.2%; Pred. No. 8.3e-11;
Matches 139; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 1576 GCTGAAGGAACTCAGTGAAGAACTGGAACTGGCAGAGAGAGGCTTGGCTTCCAAACAGCT 1635
DB 925 GCTGGAAGATCTCAACAGCAGCTCCAGAGGCGGAGGAGGCTTGGTGGCCAAACAGGA 984
QY 1636 GCAAAATGGATGAATGAAGCAAAACCAATTGCAAGCAGAGAGAGACCTGGAAACCATGAC 1695
DB 985 GGTGATCGATAAGCTGAAGGAGGAGGCGGAGCAGCAAGATTGTGATGAGACCGTTCC 1044
QY 1696 CATCTCAGGGCTCAGATGAAGTTTACTTCTGATTTTCATGCTGAAAGAGCAGCGAG 1755
DB 1045 GGTGCTGAAGGCCCGAGCGGATATCTCAAGGCGGACTTCCAGGCTGAGAGGAGGCGCG 1104
QY 1756 AGAGAAATTCATGAGGAAAGAGCAACTGGCATTGCGCTGCGCTGCTGCTGAAAGA 1815
DB 1105 GGAGAGCTGGCGAGAGAGAGAGGAGCTCTTGCAGGAGCAGCTGGAGCAGCTGCAGAGGGA 1164
QY 1816 GAA 1818
DB 1165 GTA 1167

RESULT 8
US-09-646-403-1
Sequence 1, Application US/09646403
Patent No. 6734174
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: KOVALENKO, Andrei
TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF RECEPTORS OF THE TNF/NGF RECEPTOR

TITLE OF INVENTION: AND OTHER PROTEINS
FILE REFERENCE: WALLACH-27
CURRENT APPLICATION NUMBER: US/09/646,403
CURRENT FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: IL 123758
PRIOR FILING DATE: 1998-03-19
PRIOR APPLICATION NUMBER: PCT/IL99/00158
PRIOR FILING DATE: 1999-03-18
PRIOR APPLICATION NUMBER: IL 126024
PRIOR FILING DATE: 1998-09-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 2009
TYPE: DNA
ORGANISM: Homo sapiens
US-09-646-403-1
Query Match 3.7%; Score 76.6; DB 3; Length 2009;
Best Local Similarity 57.2%; Pred. No. 8.3e-11;
Matches 139; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 1576 GCTGAAGGAACTCAGTGAAGAACTGGAACTGGCAGAGAGGCTTGGCTTCCAAACAGCT 1635
DB 930 GCTGGAAGATCTCAACAGCAGCTCCAGAGGCGGAGGAGGCTTGGTGGCCAAACAGGA 989
QY 1636 GCAAAATGGATGAATGAAGCAAAACCAATTGCAAGCAGAGAGAGACCTGGAAACCATGAC 1695
DB 990 GGTGATCGATAAGCTGAAGGAGGAGGCGGAGCAGCAAGATTGTGATGAGACCGTTCC 1049
QY 1696 CATCTCAGGGCTCAGATGAAGTTTACTTCTGATTTTCATGCTGAAAGAGCAGCGAG 1755
DB 1050 GGTGCTGAAGGCCCGAGCGGATATCTCAAGGCGGACTTCCAGGCTGAGAGGAGGCGCG 1109
QY 1756 AGAGAAATTCATGAGGAAAGAGCAACTGGCATTGCGCTGCGCTGCTGCTGAAAGA 1815
DB 1110 GGAGAGCTGGCGGAGAGAGAGGAGCTCTTGCAGGAGCAGCTGGAGCAGCTGCAGAGGGA 1169
QY 1816 GAA 1818
DB 1170 GTA 1172

RESULT 9
US-09-646-403-2
Sequence 2, Application US/09646403
Patent No. 6734174
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: KOVALENKO, Andrei
TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF RECEPTORS OF THE TNF/NGF RECEPTOR
FILE REFERENCE: WALLACH-27
CURRENT APPLICATION NUMBER: US/09/646,403
CURRENT FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: IL 123758
PRIOR FILING DATE: 1998-03-19
PRIOR APPLICATION NUMBER: PCT/IL99/00158
PRIOR FILING DATE: 1999-03-18
PRIOR APPLICATION NUMBER: IL 126024
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 2034
TYPE: DNA
ORGANISM: Homo sapiens
US-09-646-403-2
Query Match 3.7%; Score 76.6; DB 3; Length 2034;
Best Local Similarity 57.2%; Pred. No. 8.4e-11;
Matches 139; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy		1576	GCTGAAGGAACTGAGTGAAAATCTGGAACTGGCGACAGAGAGCTCTGCCTTCCAAACAGCT	1635
Db		954	GCTGGAGATCTCAAAACAGCAGCTCCACGACGCCGAGGAGCCCTGTGTGCCTCAAACAGGA	1013
Qy		1636	GCAATGGATTGAATGAAGCAAAACCATTTGCCAACGACGAAAGAGACCTTGGAAAAACATGAC	1695
Db		1014	GGTGATCGATAAGCTGAAGGAGGCGCGACGACACAAGATTGTGATGAGACCCTTC	1073
Qy		1696	CATCCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCAGCGAG	1755
Db		1074	GGTGCTGAAGGCCCCAGGCGGATATCTACAAGGCGGACTTCAGGCTGAGAGGCAGGCCCG	1133
Qy		1756	AGAGAAATTCATGAGAAAAGGACCACTGGCATTCAGCTGGCAGTTCTTGCTGAAGA	1815
Db		1134	GGAGAACTGGCCGAGAAGAGGAGCTCTTCAGGAGCAGCTGGAGCAGCTGCAGAGGA	1193
Qy		1816	GAA	1818
Db		1194	GTA	1196

RESULT 10
US-09-863-049B-3
; Sequence 3, Application US/09863049B
; Patent No. 6824972
; GENERAL INFORMATION:
; APPLICANT: Kenrick, Sue J.
; APPLICANT: Nelson, David L.
; APPLICANT: Aradhya, Swaroop
; APPLICANT: D'Urso, Michele
; APPLICANT: Woffendin, Hayley
; APPLICANT: Munnich, Arnold
; APPLICANT: Smahi, Asmae
; APPLICANT: Israel, Alain
; APPLICANT: Poustka, Annemarie
; APPLICANT: Lewis, Richard A
; APPLICANT: Levy, Moise
; APPLICANT: Heiss, Nina
; TITLE OF INVENTION: Diagnosis and Treatment of Medical Conditions Associated with Def
; FILE REFERENCE: HQ-P01961US1
; CURRENT APPLICATION NUMBER: US/09/863,049B
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/206,223
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2035
; TYPE: DNA
; ORGANISM: Human
US-09-863-049B-3

	Query Match	3.7%	Score 76.6	DB 3	Length 2035	
	Best Local Similarity	57.2%	Pred. No. 8.4e-11			
	Matches 139	Conservative 0	Mismatches 104	Indels 0	Gaps 0	
Qy	1576	GCTGAAGGAATGAGTCAAAACCTGGAACTGGCAGAGAAGCTCTGCTTCCAAAACAGCT	1635			
Db	962	GCTGGAAGATCTAAACAGCAGCTCCACGAGGCCGAGGAGCCCTGTGTGCCAAACAGGA	1021			
Qy	1636	GCAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGAGAAGACCTGGAAACCATGAC	1695			
Db	1022	GGTGATCGATAAGCTGTAAGGAGGAGGCCGAGCAGCACAAAGATTGTGATGGAGACCGTTCC	1081			
Qy	1696	CATCCTCAGGGCTCAGATGGNAAGTTACTGTTCTGTATTTTCAGTCTGAAGAAGCAGCGAG	1755			
Db	1082	GGTGCTGAAGGCCCAGGCGGATATCTACAAGGGGGACTTCAGGCTGAGAGGAGGCGGCCG	1141			
Qy	1756	AGAGAAATTCATGAGGAAAAGGAGCAACTGGCAATTGCAGCTGGCAGTTCTGCTGAAGA	1815			
Db	1142	GGAGAGCTGGCCGAGAAAGAGGAGCTCTTCGAGGAGCAGCTGAGCAGCTGCAGAGGGA	1201			

QY 1816 GAA 1818
|
Db 1202 GTA 1204

RESULT 11
US-09-543-679A-2717
; Sequence 2717, Application US/09543679A
; Patent No. 7034007
; GENERAL INFORMATION:
; APPLICANT: NYCE, Jonathan W
; TITLE OF INVENTION: LOW ADENOSINE ANTI-SENSE OLIGONUCLEOTIDE,
; COMPOSITIONS, KIT & METHOD FOR TREATMENT
; OF AIRWAY DISORDERS ASSOCIATED WITH
; BRONCHOCOINSTRUCTION, LUNG INFLAMMATION,
; NUMBER OF SEQUENCES: 3111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: EPIGENESIS PHARMACEUTICALS, INC.
; STREET: 7 Clarke Drive
; CITY: Cranbury
; STATE: NJ
; COUNTRY: USA
; ZIP: 08512
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: N/A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/543,679A
; FILING DATE: 13-Apr-2000
; CLASSIFICATION: UNKNOWN
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/127,958
; FILING DATE: 1998-08-03
; ATTORNEY/AGENT INFORMATION:
; NAME: Amzel, Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: EPI-0067191b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-409-3035
; TELEFAX: 413-254-9245
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2717:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8631 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2717
US-09-543-679A-2717

Query Match	3.7%	Score 76.6;	DB 5;	Length 8631;
Best Local Similarity	57.2%;	Pred. No. 1.9e-10;		
Matches 139;	Conservative 0;	Mismatches 104;	Indels 0;	Gaps 0;
Qy	1576	GCTGAAGGAACCTGAGTGA AAAA AACTGGA AACTGGGCAGAGAAGGCTCTGGCTTCCAAAACAGCT	1635	
Db	7562	GCTGGAAAGATCTCAACAGCAGCTCCAGCAGGCCGAGGAGGCCCTGGTGGCCAAAACAGGA	7621	
Qy	1636	GCAAAATGGATGA AATGAAGCAAAACCAATTGCCAAGCAGGAAGAGGACCTGGAAACCAATGAC	1595	
Db	7622	GGTGATCGATAAGCTGAAGGAGGAGGAGGCCGAGCAGCAACAAGATTGTGATGGAGACCGCTTCC	7681	
Qy	1696	CATCCTCAGGGCTCAGATGGAAGTTTACTGTTTCTGTATTTTCATGCTGAAAGAGCAGCGAG	1755	
Db	7682	GGTGCTGAAGGCCACGGCGGATATCTA C AAGGCCGACTCCAGGCTGAGAGGCGAGGCCG	7741	
Qy	1756	AGAGAAAATTCATGAGGAAAAGGAGCAACTGGCAITTTGCAGCTGGCAGTTCTGTGAAAGA	1815	
Db	7742	GGAGAAAGCTGGCCGAGAAAGAGAGCTCTTCGACGAGCAGCTGGAGCAGCTGCAGAGGGA	7801	
Qy	1816	GAA	1818	

Db 7802 GTA 7804

RESULT 12
US-09-767-1308
; Sequence 1308, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1308
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-767-1308

Query Match 2.4%; Score 49.8; DB 3; Length 1827;
Best Local Similarity 44.1%; Pred. No. 0.0027;
Matches 258; Conservative 0; Mismatches 322; Indels 5; Gaps 1;

QY 1104 AAAAGAAAACAAGTAATCGTTCTGAGATTGAAACCCACAGACAGAGGGGAGGACACAGAGAAAG 1163
Db 1241 AAGAGAAAAGAGAGAGATGAG 1300

QY 1164 AGAATGATGAAGAGAAAGCCCGGAGACTGTGGAAGCGAAGTGGAAAGCACTGAACTGAACTCC 1223
Db 1301 AAGAGAAAAG 1360

QY 1224 AGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGC 1283
Db 1361 AAGAGAAAAG 1420

QY 1284 TAATGATGAAGAGAAAGCCCGGAGACTGTGGAAGCGAAGTGGAAAGCACTGAACTGAACTCC 1343
Db 1421 AAGAGAAAAG 1480

QY 1344 CATCAGAGTTGATGAAACCAAGAGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGGCTTCAAG 1403
Db 1481 GAAAG 1540

QY 1404 TGAAG 1463
Db 1541 AAGAA-----GAAAG 1595

QY 1464 AATTAACTGTGTACAGATGACACACAAAGCTTCTTCAAGAACATATAATGCAATTGA 1523
Db 1596 AGAAG 1655

QY 1524 AACAATTTAGGAACTTAAAG 1583
Db 1656 AGAAG 1715

QY 1584 AACTGAGTGAAGAACTGGAAGTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1643
Db 1716 AAGAAG 1775

QY 1644 ATGAATGAAGCAACCTTTGCCAAG 1688
Db 1776 AAGAAG 1820

RESULT 13
US-09-767-16590
; Sequence 16590, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16590
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16590

Query Match 2.4%; Score 49.8; DB 3; Length 1827;
Best Local Similarity 44.1%; Pred. No. 0.0027;
Matches 258; Conservative 0; Mismatches 322; Indels 5; Gaps 1;

QY 1104 AAAAGAAAACAAGTAATCGTTCTGAGATTGAAACCCACAGACAGAGGGGAGGACACAGAGAAAG 1163
Db 1241 AAGAGAAAAGAGAGAGATGAG 1300

QY 1164 AGAATGATGAAGAGAAAGCCCGGAGACTGTGGAAGCGAAGTGGAAAGCACTGAACTGAACTCC 1223
Db 1301 AAGAGAAAAG 1360

QY 1224 AGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGC 1283
Db 1361 AAGAGAAAAG 1420

QY 1284 TAATGATGAAGAGAAAGCCCGGAGACTGTGGAAGCGAAGTGGAAAGCACTGAACTGAACTCC 1343
Db 1421 AAGAGAAAAG 1480

QY 1344 CATCAGAGTTGATGAAACCAAGAGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGGCTTCAAG 1403
Db 1481 GAAAG 1540

QY 1404 TGAAG 1463
Db 1541 AAGAA-----GAAAG 1595

QY 1464 AATTAACTGTGTACAGATGACACACAAAGCTTCTTCAAGAACATATAATGCAATTGA 1523
Db 1596 AGAAG 1655

QY 1524 AACAATTTAGGAACTTAAAG 1583
Db 1656 AGAAG 1715

QY 1584 AACTGAGTGAAGAACTGGAAGTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1643
Db 1716 AAGAAG 1775

QY 1644 ATGAATGAAGCAACCTTTGCCAAG 1688
Db 1776 AAGAAG 1820

RESULT 14
US-08-929-329-1
; Sequence 1, Application US/08929329
; Patent No. 6120770
; GENERAL INFORMATION:
; APPLICANT: Adams, John H
; APPLICANT: Dalton, John P
; APPLICANT: Kappe, Stefan
; TITLE OF INVENTION: Plasmidum Proteins Useful for Preparing
; TITLE OF INVENTION: Vaccine Compositions
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Barnes & Thornbury
; STREET: 11 S Meridian
; CITY: Indianapolis
; STATE: Indiana

```
; COUNTRY: USA
; ZIP: 46204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,329
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Breen, John P
; REGISTRATION NUMBER: 38,833
; REFERENCE/DOCKET NUMBER: 835910-28685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 231-7745
; TELEFAX: (317) 231-7433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium yoelii
; US-08-929-329-1

Query Match 2.3%; Score 46.8; DB 3; Length 5433;
Best Local Similarity 44.1%; Pred. No. 0.035;
Matches 195; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

QY 1198 AAGCGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGC 1257
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3409 ACGTGAAGAAGCTGCAAGAAATCTGAGATAATAAGAAATTTGAAGAAGCACAAGAGC 3468
QY 1258 TCATACAAACTCAGCGAAGCTGAGCTAATGAGAGAGAGACTTCAAGAAAGTGTGAGC 1317
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3469 TGGCTGGGCAAAAAGCAGAGAGAGAAAGGAAAGGCTGAAGCTGTAAAAAAGCAGA 3528
QY 1318 CCTTGAAGAGAAAATTTCTGCAATTCATCAGAGTTTGAATGAAAGCAAGAGCTTCTTTA 1377
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3529 GGAGGAAAGAAACGAATTGAAGCTGAAAGAAAGCAGAGAGGAAAGAAACGAATTGA 3588
QY 1378 TACTAACAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCATGAAATCAAAATGGAACA 1437
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3589 AGCTGAAAAGAAAGCAGAGAGAGAAAGAAACCGAATTGAAGCTGAAAAGAAAGCAGA 3648
QY 1438 GGTAAACAGAGGATGAAAAGTCCAAATTAATCTGTGTACAGATGACACACAAAGCT 1497
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3649 GGAAGAGAAAATAATTGAAGCTGCAAGAAAGCAGAGAGAGAAAGAAAGAAATTCGAAGA 3708
QY 1498 TCTTCAAGACATAATAATTCATTTGAAACAAATTTGAGGAACATAACAGAAAGAGTCAGA 1557
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3709 AGCTAAAAGAGCAGAGAGAGAAAGAAAGAAATTTGAGCTGCAAGAGAGAGAGAGA 3768
QY 1558 AAAAGTGGACAGGGCAGTGTCTGAAGGAACCTGAGTGAAGAACTGGAATCGCAGAGAGGC 1617
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3769 AAGAAAAAGGCTGAAGCTGTAAAAAAGCAGAGAGGCAAAAAAAGGCTGAAGCTGC 3828
QY 1618 TCTGGCTTCCAACAGCTGCAA 1639
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3829 AAGAAGAACCTTGAAGCTGAAA 3850

RESULT 15
US-09-902-540-1318
; Sequence 1318, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1318
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(614)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1318

Query Match 2.2%; Score 46.4; DB 3; Length 614;
Best Local Similarity 45.8%; Pred. No. 0.014;
Matches 193; Conservative 0; Mismatches 227; Indels 1; Gaps 1;

QY 1149 GCAGCACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTGTGAAGCGAAGTGG 1208
Db 30 GTAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAGGGAGGGAGAGATATAT 89
QY 1209 AAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATACAAAAC 1268
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
90 AAAAAATAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 149
QY 1269 TCAGCGAAGCTGAGCTAATGAAAGAGAGAGACTTCAAGAAAGTGTCAAGGCCCTTGAAGGA 1328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150 TGAGAAAACGAAAATATAGATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAAG 209
QY 1329 AAAATTTCTGCAATTCATCAGAGTGTGAATGAAAGCAAGAGCTTGTTTATCTAACAACA 1388
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
210 AAAAAAGAGAAAAAACAACAACAAGAAAAAGCCGAATAAAAAAACAATAACACAATA 269
QY 1389 AGTTAGAGCTACAAGTGGAAAGCA-TGCTATCAGAAATCAAAATGGAACAGGCTAAAAACA 1447
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
270 AACAACAACAAAAAGAACCAAAAAACAACAAGAAAAAAGAAAAAAGAACACAAAAA 329
QY 1448 GAGGATGAAAGTCCAAATTTAATCTGTCTACAGATGACACACAAAGCTTCTTCAAGAA 1507
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
330 ACAAAACAAAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 389
QY 1508 CATATAATATGCAATTTGAAAAACAATTTGAGGAACATAACAAAGAAAAAGAGTCAGAAAAAGTGGAC 1567
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
390 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 449
QY 1568 A 1568
Db 450 A 450

RESULT 16
US-09-949-016-3860
; Sequence 3860, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3860
;; LENGTH: 2719
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-3860

Query Match 2.2%; Score 46.2; DB 3; Length 2719;
Best Local Similarity 43.5%; Pred. No. 0.035; Mismatches 273; Indels 0; Gaps 0;
Matches 210; Conservative 0;

QY 1170 ATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAGTGGGAAGCACTGAACTCCAGGTGA 1229
DB 848 ATAAAGGTTCAAGCACTCCCAAGGCAATAGAGAACAAAGCTTTTGATCGCAATACAG 907

QY 1230 CATCTCTGTTAAGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGCTAATGA 1289
DB 908 AATCTCTCTTTGAAGAAGCTGCTTCAGCTGGCTCAGGCTTAATAGGAGATGTGGATGAAG 967

QY 1290 AGAAGAGACTTCAAGAAAGTGTGAGGCCCTTGAAGGAAATTTCTCAATTCATCAG 1349
DB 968 GAGCAGATTTACTAGGAATGGTTCGGGAAGTTGAGAACTCTATATTAGAAATACACAA 1027

QY 1350 AGTTGAATCAAAAGCAAGAGCTTGTATTACTAACAAAAAGTTAGAGCTACAAGTGGAAA 1409
DB 1028 TGTGGAACCAAAATGCTTTGACATAGTGAAGATGTTTGTAGCAAAAGTGGATG 1087

QY 1410 GCATGCTATGAAATCAAAATGGAAAGCTTAAACAGAGATGAAAAAGTCCAAATPAA 1469
DB 1088 AACTGACCTGTGAGAAAGATGTGTCAGAGGGAATTTGGAGCTGTGGAAGCAAGCCAAAC 1147

QY 1470 CTGTGCTACAGATGACACACAGACTTCTTCAAGACATAATTAATCATTTGAACAA 1529
DB 1148 TGAACATAGAGAAAGAACACAGAAATTTGGAGGAAGCTTAGGAAAGCTCGGCAGAG 1207

QY 1530 TTGAGGAAGTAAACAGAAAGAGTGCAGAAAAAGTGAGAGGCGCAGTGTGAGGAAGCTGA 1589
DB 1208 CTGAAGATGCAAGGCAAAAGCAAAAGATGACGATAGTATATCCACAGCCCAAG 1267

QY 1590 GTGAAAACTGGAAGTGGCAGAGAGGCTCTGGCTTCCAAACAGCTGCAAAATGGAATGAA 1649
DB 1268 GGAACCGGTTTACTAGAGTAGAAATGGCCCGTGTCTCATGAGCGCAACCCAGTATAAG 1327

QY 1650 TGA 1652
DB 1328 AGA 1330

RESULT 17
US-09-710-279-693
; Sequence 693, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 693
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-693

Query Match 2.2%; Score 46; DB 3; Length 627;
Best Local Similarity 46.3%; Pred. No. 0.018;
Matches 228; Conservative 0; Mismatches 255; Indels 9; Gaps 2;

QY 1071 TCAAGGAGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAATCGTTCTGAGA 1130
DB 131 TTAATCAAGTAAGTAAAAAAATAAACGAATTTAGAAGAGAAAAAGCAAAAAATTTATTTAAA 190

QY 1131 TTGAACCCAGACAGAGGGGAGACAGAGAAAGAGATGATGAAGAGAAAGCCCGGAGA 1190
DB 191 AGGTAATGATAAGATCAAGACACGCTAAAGAGCAGCTGAAGATATA-----GTTGA 245

QY 1191 CTGTTGAAGCGAAGTGGAAAGCACTGAACTCCAGGTGACATCTCTGTTTAAGGAGCTTC 1250
DB 246 AATGTTAAACAAAGACAAAAAGAAATTTGAAAAAGAGAGAGGCTCTAGATAATTTCTGA 305

QY 1251 AAGAGGCTCATACAAAAGCTCAGCGAAGCTAGCTAATGAAGAGAGACTTCAAGAAAAAGT 1310
DB 306 AAAAGCATTTAAACAAAGCAAGCAATATCTTTGAACATGTAGAAAAACA-----AAGCAA 361

QY 1311 GTCAGGCCCTTGAAGGAAAAAATTTCTCAATTTCCATCAGAGTTTCAATCAAAAGCAAGAGC 1370
DB 362 AAAAGAGTTGAAACAACCTTGATAGTCTATTAAAGAAAAATATAAATATCATATGCTT 421

QY 1371 TTGTTTATCTAAACAAAAGTTAGAGCTACAAAGTGGAAAGCATGCTATCAGAAAAATCAAAA 1430
DB 422 ATGCAAGGCTTACAAAAAGCACTTAATAGGAAAAAGAACTGTTTTCTTTATTTGAATG 481

QY 1431 TGGAAAGGCTAAACAGAGAGATGAAAAGTCCAAATTAATCTGCTACAGATGACACACA 1490
DB 482 AAGATAATCAACACAATCGGAAGTAGACGAAAAATCGAAAAATCTTTCTAAAGCATATA 541

QY 1491 ACAAGCTTCTTCAAGAACATAAATGCAATTTGAAAAACAATTTGAGAACTTAACAGAAAA 1550
DB 542 AAGAAATGAATATAAATTTAATGCTTACTCAAAAGCCATTGAGAAAGTAAAAAGAGAAA 601

QY 1551 AGTCAGAAAAAG 1562
DB 602 AACAGATGTAG 613

RESULT 18
US-09-710-279-1325
; Sequence 1325, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1325
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1325

Query Match 2.2%; Score 46; DB 3; Length 627;
Best Local Similarity 46.3%; Pred. No. 0.018;
Matches 228; Conservative 0; Mismatches 255; Indels 9; Gaps 2;

QY 1071 TCAAGGAGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAATCGTTCTGAGA 1130
DB 131 TTAATCAAGTAAGTAAAAAAATAAACGAATTTAGAAGAGAAAAAGCAAAAAATTTATTTAAA 190

QY 1131 TTGAAGCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGATGAAGAGAGGCGCGAGA 1190
Db 191 AGGTAATGATAAAGATCAAGACACACGTTAAAGAGCAGCTGAAGATATA-----GTTGA 245
QY 1191 CTGTTGAAGCGAAGTGAAGCACTGAACTCCAGGTGACATCTCTGTTTAAAGGAGCTTC 1250
Db 246 AAATGTAAACAAAGACAAAAGAAATTTGAAAAGAGAGAGAGGCTCTAGATAATTTCTGA 305
QY 1251 AAGAGGCTCATACAAAACCTCAGCGAGCTGAGCTAATGAAGAGAGACTTCAGAAAAGT 1310
Db 306 AAAAGCAATTAACAAAGCCAGCAATATCTTGAACATGTAGAAAACA-----AAGCAAGA 361
QY 1311 GTCAGGCCCTTGAAGGAAAAATTTCTGCAATTCATCAGAGTTCATGAAAGCAAGAGC 1370
Db 362 AAAAGAGTTGAACAACTTGATAGTCTATTAAAGAAAATTAATACATGATGCTT 421
QY 1371 TTGTTTATATACAAAAGTTAGAGCTACAACTGGAAGCATGCTATCAGAAATCAAAA 1430
Db 422 ATGCAAGGCTTACAAAAGAGCACTTAATAAGGAAAAGAACTGTTTCTTATTGAAATG 481
QY 1431 TGGACAGGCTAAACAGAGGATGAAAGTCCAAATTAATTAAGTGTGCTACAGATGACACACA 1490
Db 482 AAGATATGCAACACATCGGAAGTAGCGGAAAATCGAAAAGATCTTCTTAAAGCATATA 541
QY 1491 ACAAGCTTCTCAAGAACATAATAATGCATTGAAAACAAATTGAGGAACCTAACAGAAAAG 1550
Db 542 AAGAAATGAATATAATTAATGCTTACTCAAAAGCCATTGAGAAAAGTAAAGAGAAA 601
QY 1551 AGTCAGAAAAG 1562
Db 602 AACAGATGTAG 613

RESULT 19
US-09-134-001C-578
; Sequence 578, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 578
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-578

Query Match 2.2%; Score 46; DB 3; Length 633;
Best Local Similarity 46.3%; Pred. No. 0.018;
Matches 228; Conservative 0; Mismatches 255; Indels 9; Gaps 2;

QY 1071 TCAAGGAGGCCAAGAAAGAGTTTCAGATTTTGAAGAAAACAAAGTAATCGTTCTGAGA 1130
Db 137 TTAATCAAGTAAGTAAAAAATAAACGAATTAGAAGAAAAGCAAAAATTTATTTAAA 196
QY 1131 TTGAAGCCAGACAGAGGGGAGCAGAGAAAGAGATGATGAAGAAAGGCGCGAGA 1190
Db 197 AGGTAATGATAAAGATCAAGACACACGTTAAAGAGCAAGCTGAAGATATA-----GTTGA 251
QY 1191 CTGTTGAAGCGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTC 1250
Db 252 AAATGTAAACAAAGACAAAAGAAATTTGAAAAGAGAGAGGCTCTAGATAATTTCTGA 311
QY 1251 AAGAGGCTCATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAGAAAAGT 1310

Db 312 AAAAGAAATTTAAACAAGCCAGCAATATCTTGAAACATGTAGAAAACA-----AAGCAAGA 367
QY 1311 GTCAGGCCCTTGAAGGAAAAATTTGCAATTCCTCAGATTCAGAGTTGAAAGCAAGAGC 1370
Db 368 AAAAGAAAGTTGAACAACTTGATAGTCTATTAAAGAAAATATAAATACATGATGCTT 427
QY 1371 TTGTTTATACTAACAAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAAAA 1430
Db 428 ATGCAAGGCTTACAAAAGAGCACITTAATAAGGAAAAGAACTGTTTCTTATTGAAATG 487
QY 1431 TGAACAGGCTAAACAGAGGATGAAAGTCCAAATTAATTAAGTGTGCTACAGATGACACACA 1490
Db 488 AAGATAATCAACACAAATCGGAAGTAGACGGAATAATCGAAAAGATCTTCTTAAAGCATATA 547
QY 1491 ACAAGCTTCTCAAGAACATAATAATGCATTGAAAACAAATTGAGGAACCTAACAGAAAAG 1550
Db 548 AAGAAATGAATATAATTAATGCTTACTCAAAAGCCATTGAGAAAAGTAAAGAGAAA 607
QY 1551 AGTCAGAAAAG 1562
Db 608 AACAGATGTAG 619

RESULT 20
US-09-710-279-4039/c
; Sequence 4039, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4039
; LENGTH: 2999
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4039

Query Match 2.2%; Score 46; DB 3; Length 2999;
Best Local Similarity 46.3%; Pred. No. 0.042;
Matches 228; Conservative 0; Mismatches 255; Indels 9; Gaps 2;

QY 1071 TCAAGGAGGCCAAGAAAGAGTTTCAGATTTTGAAGAAAACAAAGTAATCGTTCTGAGA 1130
Db 1689 TTAATCAAGTAAGTAAAAAATAAACGAATTAGAAGAAAAGCAAAAATTTATTTAAA 1630
QY 1131 TTGAAGCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAAAGGCGCGAGA 1190
Db 1629 AGGTAATGATAAAGATCAAGACACACGTTAAAGAGCAAGCTGAAGATATA-----GTTGA 1575
QY 1191 CTGTTGAAGCGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTC 1250
Db 1574 AAATGTAAACAAAGACAAAAGAAATTTGAAAAGAGAGAGAGGCTCTAGATAATTTCTGA 1515
QY 1251 AAGAGGCTCATACAAAACCTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAGAAAAGT 1310
Db 1514 AAAGCAATTTAAACAGCCAGCAATATCTTGAACATGTAGAAAACA-----AAGCAAGA 1459
QY 1311 GTCAGGCCCTTGAAGGAAAAATTTGCAATTCCTCAGATTCAGAGTTGAAATGAAGCAAGAGC 1370
Db 1458 AAAAGAAAGTTGAACAACTTGATAGTCTATTAAAGAAAATATAAATACATGATGCTT 1399
QY 1371 TTGTTTATACTACAAAAGTTAGAGCTACAAAGTGAAGAGCATGCTATCAGAAATCAAAA 1430
Db 1398 ATCAAGAGGCTTACAAAAGAGCACTTAATAAGGAAAAGAACTGTTTCTTATTGAAATG 1399

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20343
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-20343

Query Match 2.2%; Score 45.2; DB 3; Length 705;
Best Local Similarity 45.7%; Pred. No. 0.032;
Matches 235; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 1103 GAAAGAAACAACTGATTCCTCTGAGATTGAAACCCAGACAGAGGGGACAGAGAAA 1162
DB 135 GGAAGAGACACAG 194

QY 1163 GAGAATGATGAAGAGAAAGCCCGGAGACTGTTTGGAGCGAAGTGGAACTGGAACCTC 1222
DB 195 AGAAGAGAAAGAAAG 254

QY 1223 CAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATCAAACTCAGCGAGCTGAG 1282
DB 255 GAGGAAG 314

QY 1283 CTAATGAAGAGAGACTTCAAGAAAGAGTGTGAGGCCCTTGAAGAGAAAATTCGCAATT 1342
DB 315 AGAAAAG 374

QY 1343 CCATCAGAGTTGAATGAAAGCAAGAGCTGTTTATTACTAAACAAAAGTTTAGAGCTTACAA 1402
DB 375 AG 434

QY 1403 GTGGAAGCATGCTATCAGAAATCAAAATGGAAACAGCGCTTAACACAGAGATGAAAGTCC 1462
DB 435 GAGGAG 492

QY 1463 AATTAACCTGTCTACAGATGACACACAACTCTTCTCAAGAACTATTAATGATTCG 1522
DB 493 AGAGAAAG 548

QY 1523 AAAACAATTGAGAACTAACAG 1582
DB 549 AAAAG 608

QY 1583 GAATGAGTGAATACTGGACTGGCAGAGAGG 1616
DB 609 AAAGAGACAG 642

RESULT 24
US-09-461-697-193
; Sequence 193, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Purnam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 193
; LENGTH: 696
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-461-697-193

Query Match 2.1%; Score 44.6; DB 3; Length 696;
Best Local Similarity 43.5%; Pred. No. 0.047;
Matches 203; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 1212 CACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAAACCTCA 1271
DB 171 CACAGAAAAGGGAG 230

QY 1272 CGGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAGTGTTCAGGCCCTTGAAGAGAAA 1331
DB 231 AGATCAGAAAAGAGATGAAGAAGATCAAAAACGAGAGAAAGGGAGCTGGAAAAGAGAGA 290

QY 1332 ATTCTGCAATTCATCAGAGTTGAATGAAGCAAGAGAGCTTGTATATCTACACAAAAGCT 1391
DB 291 CAAAGATGAAAAGGGGAGAGAGATGGAAAAGAGAGATAAAAATGGAATGAGAAAGGAGA 350

QY 1392 TAGAGCTACAGTGGAAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGG 1451
DB 351 AGATGCAAAAAGAGAGAGAGATGGAAAAGAGTGAAGAGGAGAGAGAGAGAGAGAGAGAG 410

QY 1452 ATGAAAAGTCCAAATTAATCTGTGTACAGATCACACACAAAGCTTCTTCAAGAACATA 1511
DB 411 TGGAAAAGAGAAAGGAGAGAGATGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 470

QY 1512 ATAATGCATTGAAAACAATTGAGGAACCTAACAGAAAAGAGTCAGAAAAGTGGACAGGG 1571
DB 471 AAAAGAGATGAAGATGGAAGAGAGAGAGAGAGATAAAAAGAGGGAAGATGTAAAAAGT 530

QY 1572 CAGTGTGAAGAACTGAGTGAAGAACTGGAAGTGGCAGAGAGGCTCTGGCTTCCAAAC 1631
DB 531 CAAAGAGATGAAGAGAGAGAGAGAGATGAAAAGAGAGATGAAGGTGGAATGAGGAAGA 590

QY 1632 AGCTGCAATGGATGAATGAAGCAAAACCATTTGCCAAGCAGGAAGAG 1678
DB 591 AGCTGMAAAGAGAAAGAGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637

RESULT 25
US-09-461-697-191
; Sequence 191, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Purnam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 191
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-191

Query Match 2.1%; Score 44.6; DB 3; Length 699;
Best Local Similarity 43.5%; Pred. No. 0.047;
Matches 203; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 1212 CACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAAACCTCA 1271
DB 174 CACAGAAAAGGGAG 233

APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: SECRETED THEM
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,228
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 3113 base pairs
Type: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-993-228-20

Query Match 2.1%; Score 44.6; DB 2; Length 3113;
Best Local Similarity 43.3%; Pred. No. 0.11;
Matches 209; Conservative 0; Mismatches 274; Indels 0; Gaps 0;
QY 1170 ATGAGAGAAAGCCCGAGACTGTGGAGCGAGTGGAGAGCACTGAACTCCAGTGA 1229
DB 139 ATAAGGTTCAAGCACTCCCAAGGCGATAGAGAACTTTTGATCGCAATACAG 198
QY 1230 CATCTCTGTTTAAGGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGCTAATGA 1289
DB 199 AATCTCTTTGAGAACTGTCTCAGCTGGCTCAGGCCTAATAGGAGATGTGGATGAAG 258
QY 1290 AGAAGAGACTTCAAGAAAGTGTGAGCCCTTGAAGGAGAAATTTCTGCAATTCCTATCAG 1349
DB 259 GAGCAGATTTTAC TAGGAATGGGTCGGAAGTTGAGAACTTATATAGAAATACACAAC 318
QY 1350 AGTTGAATGAAGCAAGAGCTGTTTATCTACTAACAAGTTAGAGCTACAGTGAAGAA 1409
DB 319 TGTGGAACCAAAATGCTTTGAACATAGTGAAGATGATTTGATAGCAAAAGTGGATG 378
QY 1410 GCATGCTATCAGAAATCAAAATGGAACAGCTTAAACAGAGGATGAAAGTCCAAATTAA 1469
DB 379 AACTGACTGTGAGAAAGATGTGCTGCAAGGGAAATTTGGAGGCTGTGAAGCAAGCCAAAC 438
QY 1470 CTGTGCTACAGATGACACAAACAAGCTTCTTCAAGAACATAATAATGCAATTTGAAACAA 1529
DB 439 TGAACCTAGAGAAAGAAACAGAGAAATTTGAGGAGAGAGCTTAGGAAGCTCGGCAGAG 498
QY 1530 TTGAGGAACTAACAGAAAGAGTCAAGAAAGTGAAGAGGAGTCTCTCAAGGAACTGA 1589
DB 499 CTGAAGATGCGGGCAAAAGCAAAAGATGACCGATGATGTATTTCCCAACAGCCACA 558
QY 1590 GTGAAAGCTGGAAGTGGCAGAGAGGCTCTGGCTTCCAAACAGCTCAATGGATGAAA 1649
DB 559 GGAACCGTTTACTAGATGAAATGGCCCGTGTCTCATGGAGCGAAACCAAGTATTAAG 618

QY 1650 TGA 1652
DB 619 AGA 621

RESULT 31

US-09-902-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 2.1%; Score 44.4; DB 3; Length 612;
Best Local Similarity 43.1%; Pred. No. 0.049;
Matches 207; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
QY 1082 AAAGAAAGAGTTTCAGATTTTGAAGAAAGAAAGAGATGATGAAGAGAAAGCCCGAGAGCTGTTGGAGC 1141
DB 26 AA 85
QY 1142 ACAGAGGGGAGCACAGAGAAAGAGATGATGAAGAGAAAGCCCGAGAGCTGTTGGAGC 1201
DB 86 AA 145
QY 1202 GAAGTGAAGAGCTGAACCTCCAGGTCACATCTCTGTTTAAAGAGCTTCAAGAGGCTCAT 1261
DB 146 AA 205
QY 1262 ACAAACCTCAGGAGAGCTGAGCTTAATGAAGAGAGAGCTTCAAGAAAGTGTGAGGCCCTT 1321
DB 206 AA 265
QY 1322 GAAGGAGAAATTTCTGCAATTTCCATCAGAGTTGAAATGAAAGCAAGAGCTTGTGTTTACT 1381
DB 266 AA 325
QY 1382 AACAAAGTTAGAGCTTACAAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGCT 1441
DB 326 AA 385
QY 1442 AAAAAAGAGAGTGAAGAGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1501
DB 386 AATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 445
QY 1502 CAGAACATAATTAATGCAATTTGAAGCAATTTGAGAGACTTACAGAGAAAGAGTCAAGAAA 1561
DB 446 AA 505

RESULT 32

US-09-248-796A-6172
; Sequence 6172, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 6172

LENGTH: 1893

TYPE: DNA

ORGANISM: Candida albicans

US-09-248-796A-6172

Query Match 2.1%; Score 44.4; DB 3; Length 1893;

Best Local Similarity 44.2%; Pred. No. 0.092;

Matches 183; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

QY 1053 AGAGCTTGAAGTGGACCTCAAGAGGCCCAAGAAAGATTTTCAGATTTTGAAGAAAA 1112

DB 1283 AGAAATTTGAATATTTTGTAGTGAAGAAAAAGAAAAAACTGAACAAAGAACTTGAAGAACTTTA 1342

QY 1113 CAAGTAATCGTCTGAGATTGAACCCAGACAGAGGGGAGCAGACAGAAAGAAATGATG 1172

DB 1343 CAAACACACGCTGAATTTGAAAAAAGTATTTGATGATAAATGCTGATTCAGTAAGA 1402

QY 1173 AAGAGAAAGCCCGGAGACTTGTGGAGCGAAGTGGAAAGCACTGAACCTCCAGGTGACAT 1232

DB 1403 CAAATTCGTATCTTGAAAAAGCCCAACAGAAAGTATTGGACCTTTGAAGCCAACTTTGAAA 1462

QY 1233 CTCGTTTAAGGAGCTTCAAGAGCTCATACAAAACCTCAGCGAGCTGAGCTTAATGAAGA 1292

DB 1463 AACTTACCAAGAACTTGATTAATAAACTGGAATTTACAAAAACATGAAAAATTTGAACA 1522

QY 1293 AGAGACTTCAAGAAAAAGTGTGAGCCCTTGAAAGGAAAAATTTCTGCAATTTCCATCAGAGT 1352

DB 1523 ATGAATTCACCAATTTGAGTAATATCATAGAGAAATTTAACTAAATTTAGACAAATTTG 1582

QY 1353 TGAATGAAAGCAAGAGCTTGTTTTATACAAAAAGTTAGAGCTACAAGTGGAAAGCA 1412

DB 1583 AAATGAAAGAAATGATCTTGTCAAATCATGAAAACTCAATGTTGAACATAAAGAAA 1642

QY 1413 TGCTATCAGAAATCAAAATGGAACAGAGCTTAAACAGAGATGAAAAGTCCAAAT 1466

DB 1643 CGTCAACAAATTTACAAATGTTTGAAGAGAAAAATAAGAAAAATCACAGCAAAAT 1696

RESULT 33

US-09-270-767-8286/c

Sequence 8286, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 8286

LENGTH: 851

TYPE: DNA

ORGANISM: Drosophila melanogaster

FEATURE:

OTHER INFORMATION: n means any nucleotide

US-09-270-767-8286

Query Match 2.1%; Score 44; DB 3; Length 851;

Best Local Similarity 44.6%; Pred. No. 0.077;

Matches 262; Conservative 0; Mismatches 320; Indels 6; Gaps 2;

QY 1031 AGGGAAGGAATCAGAAGTGGAGAGA-CTTGAAGTTGCACCTCAAGAGGCCCAAGAAAG 1089

DB 757 AGGGGAGAGAAAAGAGAGAGAGAGAAAGGAAAGGAAAGGAGGAGGAGAAAA 698

QY 1090 AGTTTCAGATTTTGAAGAAAAAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGG 1149

DB 697 AAGTAGAGAAAAAGAGAAAGAGAGAGAGAAAAAGAAAAAGAGAGAGAGAGGAA 638

QY 1150 GAGCACAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTTGGAGCGCAAGTGG 1209

DB 637 AAGAAAAAGGAAAGGGGAAAGGAAAGAGAAAAAGAAAAAGAAAGAGAGAGAGAAAA 578

QY 1210 AGCACTGAACTCCAGGTCACATCTCTGTTTAAAGAGAGCTTCAAGAGGCTTCATACAAAACT 1269

DB 577 GAAAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 518

QY 1270 CAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAAGTGTCAAGCCCTTGAAG--- 1326

DB 517 GAGGGAAGAGAAAAAGAGAGAGAGAAAAATATAGAAAAAGAGAAAAAGAGATCAAGAGGG 458

QY 1327 --GAAAAATTTCTCAATTCATCCATCAGAGTTGAATGAAAAAGCAAGAGCTTCTTTATATACTAAC 1384

DB 457 GAGAAAAAGGAGAGAGAGAAAAAGAGGGGAGAAAAAGAAAAAGAGAGAGGAGGAGAA 398

QY 1385 AAAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATCGAAACAGGCTAAA 1444

DB 397 AAGGAGAAAAAGAAAAAGAAAAATGAAAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAG 338

QY 1445 ACAGAGGATGAAAGTCCAAATTTAACTGTGCTACAGATGACACACAAAGCTTCTTCAA 1504

DB 337 AGAAAAAAGAAAGAGGGGAGGAGAAAGAGAAAGAGGAAAAAGAAAAAGAGAGAGAGAAA 278

QY 1505 GAACATAATAATGCAATTTGAAACAATTTGAGGAACTTAACAAGAAAAAGAGTCAGAAAAAGTG 1564

DB 277 AGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 218

QY 1565 GACAGGGCAGTCTGAAGGAACTGAGTGAAGAACTGGAACCTGGAACCTGCGCAGAG 1612

DB 217 GGAAGGGAGGGAG 170

RESULT 34

US-09-270-767-23568/c

Sequence 23568, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 23568

LENGTH: 851

TYPE: DNA

ORGANISM: Drosophila melanogaster

FEATURE:

OTHER INFORMATION: n means any nucleotide

US-09-270-767-23568

Query Match 2.1%; Score 44; DB 3; Length 851;

Best Local Similarity 44.6%; Pred. No. 0.077;

Matches 262; Conservative 0; Mismatches 320; Indels 6; Gaps 2;

QY 1031 AGGGAAGGAATCAGAAGTGGAGAGA-CTTGAAGTTGCACCTCAAGAGGCCCAAGAAAG 1089

DB 757 AGGGGAGAGAAAAAGAGAGAGAGAGAGAGAGAGGAAAGGAAAGGAGGAGGAGAAAA 698

QY 1090 AGTTTCAGATTTTGAAGAAAAAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGG 1149

DB 697 AAGTAGAGAAAAAGGAA 638


```
US-09-949-016-77355
; Sequence 77355, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77355
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-77355

Query Match      2.1%; Score 43.6; DB 3; Length 601;
Best Local Similarity 48.7%; Pred. No. 0.082;
Matches 115; Conservative 1; Mismatches 120; Indels 0; Gaps 0;

QY 1328 AAAAAATTCGCAATCCATCAGAGTTGAATGAAAAAGCAAGAGCTGTTGTTATATACTAACAAA 1387
DB 130 AAGATATTTTGAAGTGTATATGAAGTGTATTAAGAGAAAAAGAGGAATTCAGAGAAAAAG 189

QY 1388 AAGTTAGAGCTACAGTGGAAAGCATCTATCAGAAATCAAAATGGAACAGGCTTAAACA 1447
DB 190 ACACCAAGAAAAAAGACAGGCATAATTTAAAGCTTCTGAAGAACTTCTAGAAAAATA 249

QY 1448 GAGGATGAAAGTCCAAATTAATCTGCTACAGATGACACACAAACAGCTTCTTCAAGAA 1507
DB 250 GAACATTAAATTCAAACATAAATGGTCTAACCGTTAAACAGCAGACTGAAATGAATTGA 309

QY 1508 CATAATAATGCATGAAAAACAATTTAGGAACTAACAGAAAAAGAGTCAGAAAAAGT 1563
DB 310 CAGGAATATTAACTACGAGAAATGAGCAACAGACGACACAAAAAAGACAGAAAAAT 365

RESULT 38
US-09-949-016-16602/c
; Sequence 16602, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16602
; LENGTH: 55841
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16602

Query Match      2.1%; Score 43.6; DB 3; Length 55841;
Best Local Similarity 46.0%; Pred. No. 0.99;
Matches 148; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
```

```
QY 1276 AGCTCAGCTAATGAAGAAGAGAGCTTCAAGAAAAAGTGTGAGGCCCTTTGAAAGGAAAAATTC 1335
DB 31546 AACCTAGCGACAGAATGAGACTCTGTCTTAAAAAAGAGAGAGAGAGAGAGAAAAAGA 31487

QY 1336 TGCAATTCATCAGAGTTGAAATGAAAAAGCAAGAGCTGTTTATATACTAAACAAAAAGTTAGA 1395
DB 31486 AAGAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 31427

QY 1396 GCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACACAGAGATGA 1455
DB 31426 AAGNAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 31367

QY 1456 AAAAGTCCAAATTAATCTGCTACAGATGACACACAAACAAAGCTTCTTCAAGAAACATATAA 1515
DB 31366 AAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 31307

QY 1516 TGCATTGAAAAACAATTGAGGAACCTAACCAAGAAAAAGAGTCAGAAAAAGTGGACAGGCGACT 1575
DB 31306 AAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 31247

QY 1576 GCTGAAGGAAGTCAAGTGAAGAAA 1597
DB 31246 AAGAAAAAGAAAGAAAGAAAGAAA 31225

RESULT 39
US-09-336-115C-23
; Sequence 23, Application US/09336115C
; Patent No. 6576244
; GENERAL INFORMATION:
; APPLICANT: Weltzin, Richard A.
; APPLICANT: Guy, Bruno
; TITLE OF INVENTION: LT and CT in Parenteral Immunization
; TITLE OF INVENTION: Methods Against Helicobacter Infection
; FILE REFERENCE: 06132/055002
; CURRENT APPLICATION NUMBER: US/09/336,115C
; CURRENT FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: US 09/100,258
; PRIOR FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)...(1002)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (106)...(166)
US-09-336-115C-23
```

```
Query Match      2.1%; Score 43.4; DB 3; Length 1149;
Best Local Similarity 46.3%; Pred. No. 0.13;
Matches 143; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 1081 CAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAGATTAATCGTTCTGAGATTGAACCCCA 1140
DB 264 CACTAAAAGCGATTTTGACATGATTAAGCAACGAAATCCTAATTTTGTATTTTGACAAGCT 323

QY 1141 GACAGAGGGGACACAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTTGAAG 1200
DB 324 TAAAGAGAAAGAAAGAAAGAGCGCTTGATGATCAGCTATTCGACCGCCCTTGTAAGAAA 393

QY 1201 CGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCA 1260
DB 384 TGAAGCTAAAACCGAGAAATTTGACAGCACTCCAGAAATTTAAAGCGATGATGAAGCGGT 443

QY 1261 TACAAACTCAGGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGTGTCAGGCCCT 1320
DB 444 TAAAAACAGGCTTTAGTGAATTTTGGCTAAAAAAGGCTGAAGAAAGTGAAGAAAGT 503
```

QY 1321 TGAAGGAAAAATTCGCAATTCATCCATCAGAGTTGAAATGAAAGCAAGAGCTGTGTTATAC 1380
Db 504 CCAATCCAGAAAAAGAAATGCAAGATTTTACAACGCTAAACAAGATCAGCTTTTGT 563
QY 1381 TAACAAAA 1389
Db 564 CAAGCAGA 572

```

RESULT 40
US-09-710-279-4002/c
; Sequence 4002, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMWELLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4002
; LENGTH: 3305
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4002

```

```

RESULT 41
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22

```


Db 211 GATCGTGGAGAAACACAAAGGAAGAGCCCAAGGTTGAGGCGAGAGAAAAAGGCGAGCAA 270
Qy 1157 GAGAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAGTGAAGCACTG 1216
Db 271 GAGCGAGAAAGAGCGAGAGAAAGGCGAGAGAAAGGCGAGAGAGCGAGAGAA 330
Qy 1217 AACCTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAAACTCAGCGA 1275
Db 331 GAGGCTCCGCTGCTGAGTCCGCAACAGAGACTACACACATTTGCTACTCTCTCATCGA 389

RESULT 45

US-10-104-047-1957
; Sequence 1957, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1957
; LENGTH: 1685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1957

Query Match 2.1%; Score 42.8; DB 3; Length 1685;
Best Local Similarity 44.2%; Pred. No. 0.24;
Matches 176; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

Qy 1281 AGCTAATGAAGAGAGACTTCAAGAAAGTGTACAGCCCTTGAAGGAAATTTCTGCAA 1340
Db 721 AGAAGATGAAGAGAGACTTCAAGAAAGTGTACAGCCCTTGAAGGAAATTTCTGCAA 780
Qy 1341 TTCCATCAGAGTTGAATGAAGAGAGAGCTTTTATCTACTAACAAGAGTTAGAGCTAC 1400
Db 781 AAAAGGGAAGAGAGATGGAAGAGAGAGATGAAGAGAGAGAGAGAGAGAGATGCAAA 840
Qy 1401 AAGTGGAAAGAGATGCTATCAGAAATCAAAATGGAACAGGCTTAAACAGAGGATGAAAGT 1460
Db 841 AGAGAAAGAGATGGAAG 900
Qy 1461 CCAATTTAACTGTGCTACAGATGACACACAAAGCTTCTTCAAGAACATAATAATGCAT 1520
Db 901 GAAAGGAGAGATGAAAG 960
Qy 1521 TGAAGCAATTTGAGGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1580
Db 961 TGAAGATGGAAG 1020
Qy 1581 AGCAACTGAGTGAAG 1640
Db 1021 TGAAG 1080
Qy 1641 TGGATGAATGAAGCAACCACTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1678
Db 1081 AGAGAAAG 1118

RESULT 46

US-09-614-221A-426
; Sequence 426, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulolini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614.221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 426
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-426

Query Match 2.1%; Score 42.6; DB 3; Length 345;
Best Local Similarity 47.5%; Pred. No. 0.12;
Matches 126; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy 1435 ACAGGCTAAACAGAGAGATGAAAGTCCAAATTAACCTGTGCTACAGATGACACACAA 1494
Db 33 ACAGCTGAAAGGAAGCCACGAAATAGTATCAAGCTAGAAAGTACAGACAGATAA 92
Qy 1495 GCTTCTTCAAGACATAATATGCAATGAAAACAAATTCAGGAACCTAACAGAAAAGATC 1554
Db 93 GTTGAAGCAAGCCAGAGACTGATGCGCAAGGAAATCGACTCATACAAAATTCAAAAAGA 152
Qy 1555 AGAAAAAGTGACAGAGGCGAGTGTGAAGAGACTGAGTGAAAACTGGAACTGGCAGAGAA 1614
Db 153 CAGGAATTTGAAGAGATTTGAACAAAAGAAATGCCGGTGTGTTGGTGAATTTGAAAAAGAA 212
Qy 1615 GGCTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCATTTGCCAAGCAGGA 1674
Db 213 AGCAGAGGCTGTGTCGCAAGGTGAATAGCTGAGATTAGAATAATTCAGAGAGAGAAAAA 272
Qy 1675 AGAGGAGCTGGAACCATGACCATC 1699
Db 273 GGATGAGCTTGTCAAAAATTTTGATC 297

RESULT 47

US-09-902-540-1280
; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1039)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

Query Match 2.1%; Score 42.6; DB 3; Length 1039;
Best Local Similarity 51.0%; Pred. No. 0.21;
Matches 99; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1375 TTATCTAAACAAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGA 1434
Db 380 TTAAAAATAAAAAATTAATTTACAAACAAAAATAAAAGAAAAAATAAAAAATAAATA 439

```
QY 1435 ACAGGCTAAACAGAGGATGAAAAGTCCAAATTAACTGTCTACAGATGACACACAA 1494
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 440 ATATGTAAATAAATAACACACAAACACCCAAAAAATAAATAAATAAATAAATAA 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1495 GCTTCTTCAAGAACATAATAATGATGAAACAAATTGAGGAAGTCAACAGAAAGAGTC 1554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 ATGAAACACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 559
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1555 AGAAAAAGTGGACA 1568
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 AAAAAAGATAAAAA 573
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 48
US-09-949-016-14000
; Sequence 14000, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14000
; LENGTH: 154626
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(154626)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14000
```

```
Query Match 2.0%; Score 42.4; DB 3; Length 154626;
Best Local Similarity 48.7%; Pred. No. 3.8;
Matches 115; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 1328 AAAAAATCTGCAATTCATCAGAGTTGAATGAAAAGCAGAGCTTGTATTATCTAACAAA 1387
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 58226 AAGATATTTGTAAGTGTATATGAAGTGTATAAAGAGAAAAGAGGAATTTGAAGGAAAAAG 58285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1388 AAGTTAGAGCTACAAGTGGAAAGCAGCTATCAGAAATCAAAATGGAACAGGCTAAAAACA 1447
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 58286 ACACCAAGAAAAAAGACAGCGCATAATTTAAAGCTTCTGAAGAACTTCTAGAAATATA 58345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1448 GAGGATGAAAAGTCCAAATTAATCTGTCTACAGATGACACACAAACAGCTTCTTCAAGAA 1507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 58346 GAACATTAAATTCACACATAAATGGTCTAACGGTTAAACAGCAGACTGAATTTGAATGA 58405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1508 CATAATAATGCAATGAAAAACAATTTGAGGAAGTCAACAGAAAAAGAGTCAGAAAAAGT 1563
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 58406 CAGGAATATTAACTACCAAGAAATGAGCAACAGAACGACAAAAAAGACAGAAAAAT 58461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 49
US-09-640-173-53/c
; Sequence 53, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; METHODS OF USE THEREFOR
; FILE REFERENCE: 210121.484C2

QY 1333 TTCTGCAATTCATCAGAGTTGAATGAAAAGCAGAGCTTGTATTATCTAACAAAAAGTT 1392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TTCTTTTCTTTTNTGAATTAANAANAAGNAAAAAANAANAANAANAANAANAANAANA 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1393 AGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGGA 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1453 TGAAGAAGTCCAATTAATCTGTCTACAGATGACACAAACAAGCTTCTTCAAGAACATAA 1512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
; CURRENT APPLICATION NUMBER: US/09/640,173
; CURRENT FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; TYPE: DNA
; LENGTH: 396
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

Query Match 2.0%; Score 41.8; DB 3; Length 396;
Best Local Similarity 47.6%; Pred. No. 0.21;
Matches 109; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 1333 TTCTGCAATTCATCAGAGTTGAATGAAAAGCAGAGCTTGTATTATCTAACAAAAAGTT 1392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TTCTTTTCTTTTNTGAATTAANAANAAGNAAAAAANAANAANAANAANAANAANAANA 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1393 AGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGGA 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1453 TGAAGAAGTCCAATTAATCTGTCTACAGATGACACAAACAAGCTTCTTCAAGAACATAA 1512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 151 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 92
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1513 TAATGCATTGAAAACAATTTGAGGAAGTCAACAGAAAGAGTCAGAAAAA 1561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 43
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 50
US-09-713-550-53/c
; Sequence 53, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713,550
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53
```

```
Query Match 2.0%; Score 41.8; DB 3; Length 396;
Best Local Similarity 47.6%; Pred. No. 0.21;
Matches 109; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 1333 TTCTGCAATTCATCAGAGTTGAATGAAAAGCAGAGCTTGTATTATCTAACAAAAAGTT 1392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 TTCTTTTCTTTTNTGAATTAANAANAAGNAAAAAANAANAANAANAANAANAANAANA 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1393 AGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGGA 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1453 TGAAGAAGTCCAATTAATCTGTCTACAGATGACACAAACAAGCTTCTTCAAGAACATAA 1512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Db 151 AA 92
Qy 1513 TAATGCATTGAAAAACAATTGAGGAACTAACAAAGHAAAGAGTCAGAAAAA 1561
Db 91 AA 43

Search completed: May 29, 2006, 11:34:03
Job time : 673 secs

THIS PAGE BLANK (USPTO)


```
C 91 50.2 2.4 803 8 US-10-437-963-72176
C 92 50 2.4 1243 9 US-10-425-115-172717
C 93 49.8 2.4 600 10 US-10-972-079-13288
C 94 49.4 2.4 1155 8 US-10-221-714A-475
C 95 49.2 2.4 446 3 US-09-960-352-3400
C 96 49.2 2.4 927 9 US-10-425-115-126557
C 97 49.2 2.4 1297 9 US-10-425-115-67687
C 98 48.8 2.3 629 7 US-10-125-968-78
C 99 48.4 2.3 5142 6 US-10-097-340-42
C 100 48.4 2.3 5142 13 US-11-050-926-42
C 101 48 2.3 344 3 US-09-960-352-1036
C 102 48 2.3 447 8 US-10-021-323-9573
C 103 47.8 2.3 14006 7 US-10-311-455-1931
C 104 47.6 2.3 516 3 US-09-960-352-5785
C 105 47.4 2.3 6486 8 US-10-433-793-80
C 106 47.4 2.3 6668 7 US-10-311-453-1670
C 107 47.2 2.3 4668 9 US-10-674-124A-12575
C 108 47.2 2.3 1214 7 US-10-424-599-102083
C 109 46.8 2.3 3913 7 US-10-233-045-12
C 110 46.6 2.2 6171 7 US-10-311-455-761
C 111 46.6 2.2 7571 7 US-10-311-455-500
C 112 46.4 2.2 767 8 US-10-424-599-70401
C 113 46.4 2.2 1029 9 US-10-425-115-123086
C 114 46.4 2.2 2418 7 US-10-369-493-23839
C 115 46.4 2.2 9539 6 US-10-239-676-52
C 116 46.4 2.2 9539 7 US-10-240-453-54
C 117 46.2 2.2 411 8 US-10-424-599-79490
C 118 46.2 2.2 899 7 US-10-369-493-29260
C 119 46.2 2.2 921 9 US-10-425-115-38710
C 120 46.2 2.2 1204 7 US-10-437-963-77858
C 121 46.2 2.2 2523 7 US-10-205-219-158
C 122 46.2 2.2 3791 7 US-10-240-965-3
C 123 46.2 2.2 3825 7 US-10-422-522-8
C 124 46.2 2.2 4286 6 US-10-044-090-549
C 125 46.2 2.2 6145 7 US-10-311-455-945
C 126 46.2 2.2 3673778 7 US-10-312-841-2
C 127 46.2 2.2 627 10 US-10-793-626-693
C 128 46 2.2 627 10 US-10-793-626-1325
C 129 46 2.2 633 8 US-10-724-972A-1807
C 130 46 2.2 2999 10 US-10-793-626-4039
C 131 46 2.2 3691 10 US-10-793-626-3338
C 132 46 2.2 10170 16 US-11-136-527-1817
C 133 46 2.2 5625 9 US-10-723-860-7708
C 134 45.8 2.2 2350 9 US-10-602-494-302
C 135 45.6 2.2 5750 8 US-10-221-714A-433
C 136 45.6 2.2 6301 7 US-10-311-455-26
C 137 45.6 2.2 484 9 US-10-425-115-172345
C 138 45.4 2.2 809 9 US-10-363-345A-20409
C 139 45.4 2.2 809 9 US-10-363-345A-20410
C 140 45.4 2.2 809 10 US-10-363-483A-20409
C 141 45.4 2.2 809 10 US-10-363-483A-20410
C 142 45.4 2.2 868 8 US-10-437-963-42302
C 143 45.4 2.2 1016 8 US-10-424-599-1347
C 144 45.4 2.2 6010 9 US-10-723-860-57
C 145 45.4 2.2 6010 9 US-10-723-860-105
C 146 45.4 2.2 6010 9 US-10-723-860-4858
C 147 45.4 2.2 7442 8 US-10-221-714A-409
C 148 45.4 2.2 17934 7 US-10-311-455-1692
C 149 45.4 2.2 17934 7 US-10-311-455-1692
C 150 45.4 2.2 17934 7 US-10-311-455-1692
```

ALIGNMENTS

```
RESULT 1
US-10-609-133-1
; Sequence 1, Application US/10609133
; Publication No. US20040191798A1
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut Health Center
; APPLICANT: Sarfarazi, Mansoor
; APPLICANT: Rezaie, Tayebbeh
```

```
; APPLICANT: Child, Anne H
; TITLE OF INVENTION: Optineurin and Glaucoma
; FILE REFERENCE: UCT-0046P2
; CURRENT APPLICATION NUMBER: US/10/609,133
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: 10,281,457
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 10,090,118
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 10,060,981
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: 60/344,754
; PRIOR FILING DATE: 2001-12-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2077
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; AUTHORS: Rezaie, T., Child, A., Hitchings, R., Brice, G., Miller, L.,
; AUTHORS: Coca-Prodos, M., Heon, E., Krupin, T., Ritch, R., Kreutzer, D.,
; AUTHORS: Crick, R.P., and Sarfarazi, M.
; TITLE: Adult-Onset Primary Open-Angle Glaucoma Caused by Mutations in
; TITLE: Optineurin
; JOURNAL: Science
; VOLUME: 295
; ISSUE: 5557
; PAGES: 1077-1079
; DATE: 2002
; DATABASE ACCESSION NUMBER: AF420371
; DATABASE ENTRY DATE: 2002-02-11
; RELEVANT RESIDUES: (1)..(2077)
; US-10-609-133-1

Query Match 99.9%; Score 2075.4; DB 9; Length 2077;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCCGGTGGAGTGTCTCTCAGGGGGGACATGCGGAGGAAACAGTACCTTGAGCGA 60
Db 1 ATCCCGGTGGAGTGTCTCTCAGGGGGGACATGCGGAGGAAACAGTACCTTGAGCGA 60
QY 61 AGCCAAAGCCGGCGGAGGTGGCTTTCATAGTGGTGGTGCACCTTCTGCGCTTGA 120
Db 61 AGCCAAAGCCGGCGGAGGTGGCTTTCATAGTGGTGGTGCACCTTCTGCGCTTGA 120
QY 121 TGAGCGGTACGCTCTGTAAACCCCACTTCTCAGCTTGAACAGCTGCTGGTTTCA 180
Db 121 TGAGCGGTACGCTCTGTAAACCCCACTTCTCAGCTTGAACAGCTGCTGGTTTCA 180
QY 181 ATTAATGAAGATTAGTCAGTGACAGGCTTGGTGGTGGTGGTGGTGGTGGTGGT 240
Db 181 ATTAATGAAGATTAGTCAGTGACAGGCTTGGTGGTGGTGGTGGTGGTGGTGGT 240
QY 241 AAATGTCAAATGTAACTGGAGAGAAAGTGGGCACTTTTGGAGTGACTTTTCCAC 300
Db 241 AAATGTCAAATGTAACTGGAGAGAAAGTGGGCACTTTTGGAGTGACTTTTCCAC 300
QY 301 AACTTTTGAATGTCCCATCAACCTTCAAGTGCCTCACTGAAAGAGGACAGCCCC 360
Db 301 AACTTTTGAATGTCCCATCAACCTTCAAGTGCCTCACTGAAAGAGGACAGCCCC 360
QY 361 TGAAGCAGCAAGAAATGGACCCCGCCACCTGCGCCACCAACCTGGACAGTTTACCCC 420
Db 361 TGAAGCAGCAAGAAATGGACCCCGCCACCTGCGCCACCAACCTGGACAGTTTACCCC 420
QY 421 GGAGAGAGCTGCTGCAGCAGATGAAAGAGCTCTCTGACCAAGAACACAGCTGAAG 480
Db 421 GGAGAGAGCTGCTGCAGCAGATGAAAGAGCTCTCTGACCAAGAACACAGCTGAAG 480
QY 481 CATGAAGCTTAATTAATCAAGCCATGAAGGAGATTGAGGAGCTTTCGCGCTTGAC 540
Db 481 CATGAAGCTTAATTAATCAAGCCATGAAGGAGATTGAGGAGCTTTCGCGCTTGAC 540
```


Db 76 AGCCAAGCCGGCGGCGAGGTGTGGCTTTGATAGCTGGTGGTGCACCTTCCTGGCCCTTGGA 135
Qy 121 TGAGCGGTACGGCTCTGTAAACCCAACTTCCTCACCTTTGAAACAGCTGCCTGTTTCAGC 180
Db 136 TGAGCGGTACGGCTCTGTAAACCCAACTTCCTCACCTTTGAAACAGCTGCCTGTTTCAGC 195
Qy 181 ATTTAATGAAGATTAGTCACTGACAGGCTGGTGTGCTGAGTCCGACATAGAGAATCAA 240
Db 196 ATTTAATGAAGATTAGTCACTGACAGGCTGGTGTGCTGAGTCCGACATAGAGAATCAA 255
Qy 241 AAATGTCAAAATGTAACTGAGAGAAAGTGGGCAACTTTT-GGAGTGACTTTTCACAG 299
Db 256 AAATGTCAAAATGTAACTGAGAGAAAGTGGGCAACTTTTGGGAGTGACTTTTCACAG 315
Qy 300 GAACTTCTGCAATGTCCTCACTCACTCTCAGCTGCTCACTGAAAGGAGGACAGCCCA 359
Db 316 GAACTTCTGCAATGTCCTCACTCACTCTCAGCTGCTCACTGAAAGGAGGACAGCCCA 375
Qy 360 GTGAAGACACAGAAATGGACCCGCCCACTGGCCCAACCCAACTGGACAGCTTTTACCC 419
Db 376 GTGAAGACACAGAAATGACCCGCCCACTGGCCCAACCCAACTGGACAGCTTTTACCC 435
Qy 420 CGGAGGAGCTGTGACAGAGATGAAGAGCTCTGACCAAGAAACCACTGAGTGAAGAAAG 479
Db 436 CGGAGGAGCTGTGACAGAGATGAAGAGCTCTGACCGAGAAACCACTGAGTGAAGAAAG 495
Qy 480 CCATGAAGCTTAATTAATCAAGCCATCAAGGGAGATTTGAGGAGCTTTTGGCCCTGGACAG 539
Db 496 CCATGAAGCTTAATTAATCAAGCCATCAAGGGAGATTTGAGGAGCTTTTGGCCCTGGACAG 555
Qy 540 AGAAAACAGAAAGAAAGCCAGTTTTTTTGAAGATACAGAGCAAGAAAGAAAGAGCGTC 599
Db 556 AGAAACAGAAAGAAAGCCAGTTTTTTTGAAGATACAGAGCAAGAAAGAAAGAGCGTC 615
Qy 600 TAATGGCTTGAAGTATGAGAAATGAGAAATGAGAAAGAGCTTGAAAACTTAAAGGGA 659
Db 616 TAATGGCTTGAAGTATGAGAAATGAGAAATGAGAAAGAGCTTGAAAACTTAAAGGGA 675
Qy 660 AATCAGAAAGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAGCGG 719
Db 676 AATCAGAAAGTTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAGCGG 735
Qy 720 AGCAGAAAAGGACCAAGCTCAGACCCAGGTGGTGAAGCTTACAGCAGAGAAAGCAGACC 779
Db 736 AGCAGAAAAGGACCAAGCTCAGACCCAGGTGGTGAAGCTTACAGCAGAGAAAGCAGACC 795
Qy 780 TGTGGGCATCGTGTCTGAACTGACGCTCAAGCTCAAGCTCAAGCGGCTCCTCAGAAAGATT 839
Db 796 TGTGGGCATCGTGTCTGAACTGACGCTCAAGCTCAAGCTCAAGCGGCTCCTCAGAAAGATT 855
Qy 840 CCTTTGTTGAAATTAGATGGCTGAGGAGCCAGCAGAGGCTCAGTAAAGAAATCAAGC 899
Db 856 CCTTTGTTGAAATTAGATGGCTGAGGAGCCAGCAGAGGCTCAGTAAAGAAATCAAGC 915
Qy 900 ATAGTCTGGGCCACGAGAACAGTCTCCACTGGCAGCGCATTTGCTAAATATAGAGCA 959
Db 916 ATAGTCTGGGTCCACAGAGAACAGTCTCCACTGGCAGCGCATTTGCTACTATAGAGGA 975
Qy 960 GATCTCAGATGGGCGAAGAAATTAATTGAAACATGAGGAGTTAACTGTGAGCCAGCTCC 1019
Db 976 GATCTCAGATGGGCGAAGAAATTAATTGAAACATGAGGAGTTAACTGTGAGCCAGCTCC 1035
Qy 1020 TGCCTGTGCTTAGGGAAGGGAATCAGAAAGTGGAGAGCTTGAAGTGGCACTCAAGGAGG 1079
Db 1036 TGCCTGTGCTTAGGGAAGGGAATCAGAAAGTGGAGAGCTTGAAGTGGCACTCAAGGAGG 1095
Qy 1080 CCAAGAAAAGAGTTTCAGATTTTGAAGAAAACAAAGTAATCGTTCTGAGATTGAAACCC 1139
Db 1096 CCAAGAAAAGAGTTTCAGATTTTGAAGAAAACAAAGTAATCGTTCTGAGATTGAAACCC 1155
Qy 1140 AGACAGAGGGAGCAGAGAGAAAGAAATGATGAAGAAAGGCCCCGGAGACTGTTGGAA 1199
Db 1156 AGACAGAGGGAGCAGAGAGAAAGAAATGATGAAGAAAGGCCCCGGAGACTGTTGGAA 1215

Qy 1200 GCGAAGTGGAAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTC 1259
Db 1216 GCGAAGTGGAAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTC 1275
Qy 1260 ATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAAGAGACTTCAAGAAAAAGTGTCAAGCCC 1319
Db 1276 ATACAAAACCTCAGCGAAGCTGAGCTTAATGAAGAAGAGACTTCAAGAAAAAGTGTCAAGCCC 1335
Qy 1320 TTGAAGGAAAAAATTTCTGCATTTCCATCAGAGTTCGATGAAAGCAAGAGCTTGTTTATA 1379
Db 1336 TTGAAGGAAAAAATTTCTGCATTTCCATCAGAGTTCGATGAAAGCAAGAGCTTGTTTATC 1395
Qy 1380 CTAAACAAAAAGTTAGAGCTTACAAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAACAGG 1439
Db 1396 CTAAACAAAAAGTTAGAGCTTACAAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAACAGG 1455
Qy 1440 CTAAACAAAAAGTGAAGATGAAAGTCCAAATTAATCTGTCTCAGATGACACAAAGAGCTTC 1499
Db 1456 CTAAACAAAAAGTGAAGATGAAAGTCCAAATTAATCTGTCTCAGATGACACAAAGAGCTTC 1515
Qy 1500 TTCAAGAACATTAATTAATGCAATTTGAAACCAATTTGAGGAACTAACAGAAAGAGTCAAGAA 1559
Db 1516 TTCAAGAACATTAATTAATGCAATTTGAAACCAATTTGAGGAACTAACAGAAAGAGTCAAGAA 1575
Qy 1560 AAGTGGACAGGGCAGTGTGTAAGGAACTGAGTGAAGAACTGGCACTGGCAGAGAAAGCTC 1619
Db 1576 AAGTGGACAGGGCAGTGTGTAAGGAACTGAGTGAAGAACTGGCACTGGCAGAGAAAGCTC 1635
Qy 1620 TGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCAATTTGCCAAGCGAAAGAGG 1679
Db 1636 TGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAAAACCAATTTGCCAAGCGAAAGAGG 1695
Qy 1680 ACTGGAAACATGACCATCTCAGGCTCAGATGGAAGTCTTACTGTTCTGATTTTCATG 1739
Db 1696 ACTGGAAACATGACCATCTCAGGCTCAGATGGAAGTCTTACTGTTCTGATTTTCATG 1755
Qy 1740 CTGAAGAGCAGCAGAGAGAGAAATTCATGAGAAAGAGCAAACTGGCACTTCGAGCTGG 1799
Db 1756 CTGAAGAGCAGCAGAGAGAGAAATTCATGAGAAAGAGCAAACTGGCACTTCGAGCTGG 1815
Qy 1800 CAGTCTCTGCTGAAAGAGAAATGATGTTTTCGAAAGACGAGCGAGGCAAGTCTTGATGGAGA 1859
Db 1816 CAGTCTCTGCTGAAAGAGAAATGATGTTTTCGAAAGACGAGCGAGGCAAGTCTTGATGGAGA 1875
Qy 1860 TGCAGAGTGTCTCAGGCGGAGAACAAAGTACTCTGACACAGAGGCTTACCTTGTTCAAA 1919
Db 1876 TGCAGAGTGTCTCAGGCGGAGAACAAAGTACTCTGACACAGAGGCTTACCTTGTTCAAA 1935
Qy 1920 GAGGAGCTGAGGACAGGGAAGTGGCGGCAACAGCGGAATATTCGATTCATTCCTGCCCA 1979
Db 1936 GAGGAGCTGAGGACAGGGAAGTGGCGGCAACAGCGGAATATTCGATTCATTCCTGCCCA 1995
Qy 1980 AGTGGAGAGGTTCTGCTGACATAGACACGTTACAGATTACGTTGATGGATTCATCA 2039
Db 1996 AGTGGAGAGGTTCTGCTGACATAGACACGTTACAGATTACGTTGATGGATTCATCA 2055
Qy 2040 TTTAAGTGTGATGATATCACCTCCCAAACTGTTGGT 2077
Db 2056 TTTAAGTGTGATGATATCACCTCCCAAACTGTTGGT 2093

RESULT 3

US-10-342-887-128
; Sequence 128, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna

; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 128
; LENGTH: 2139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-128

Query Match 99.0%; Score 2056.4; DB 8; Length 2139;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2071; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY	1	ATCCCGTGGGAGTCTCTCCAGCGGCGACGATGCCAGGAAACAGTGAACCTTGACCGA	60
DB	16	ATCCCGTGGGAGTCTCTCCAGCGGCGACGATGCCAGGAAACAGTGAACCTTGACCGA	75
QY	61	AGCCAAAGCCGGCGGAGGTGGCTTTGATAGCTGGTGGTGCACCTTCCTGGCTTGA	120
DB	76	AGCCAAAGCCGGCGGAGGTGGCTTTGATAGCTGGTGGTGCACCTTCCTGGCTTGA	135
QY	121	TGAGCGGTACGCTCTGTAAACCAACTCTCTCACTTGAAGAGAGTGGTGGTTCAGC	180
DB	136	TGAGCGGTACGCTCTGTAAACCAACTCTCTCACTTGAAGAGAGTGGTGGTTCAGC	195
QY	181	ATTAAATGAAGATTAGTCAGTGAAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT	240
DB	196	ATTAAATGAAGATTAGTCAGTGAAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT	255
QY	241	AAATGTCCAAATGTAATCGAGAGAAAGTGGGCACTTTTGGAGTGGTGGTGGTGGT	299
DB	256	AAATGTCCAAATGTAATCGAGAGAAAGTGGGCACTTTTGGAGTGGTGGTGGTGGT	315
QY	300	GAACTTCTGCAATGTCCCACTCAAGCTGCTCACTGAGTGGTGGTGGTGGTGGTGGT	359
DB	316	GAACTTCTGCAATGTCCCACTCAAGCTGCTCACTGAGTGGTGGTGGTGGTGGTGGT	375
QY	360	GTAAAGCACAGGAAATGGAACCCCGACCTGGCCCAACCCCAACCTGGACACCTTACC	419
DB	376	GTAAAGCACAGGAAATGGAACCCCGACCTGGCCCAACCCCAACCTGGACACCTTACC	435
QY	420	CGGAGGAGTGTGCGAGCAGATGAAAGAGCTCTGACCGAGAACCCAGCTGAAAGAG	479
DB	436	CGGAGGAGTGTGCGAGCAGATGAAAGAGCTCTGACCGAGAACCCAGCTGAAAGAG	495
QY	480	CCATGAAGCTAAATATCAAGCCTGAAAGGAGATTTGAGGAGCTTTGGGCTGGACAG	539
DB	496	CCATGAAGCTAAATATCAAGCCTGAAAGGAGATTTGAGGAGCTTTGGGCTGGACAG	555
QY	540	AGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	599
DB	556	AGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	615
QY	600	TAATGGCTTGAATATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT	659
DB	616	TAATGGCTTGAATATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT	675
QY	660	AATCAGAAAGGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	719
DB	676	AATCAGAAAGGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	735
QY	720	AGCAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	779

Db	736	AGCAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	795
QY	780	TGTTGGGCATCGTGTCTGAACCTCAGCTCAAGCTGAACCTCAGCGGCTCTCTCAGAA	839
Db	796	TGTTGGGCATCGTGTCTGAACCTCAGCTCAAGCTGAACCTCAGCGGCTCTCTCAGAA	855
QY	840	CCTTTGTGTAATTAGGATGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	899
Db	856	CCTTTGTGTAATTAGGATGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	915
QY	900	ATAGTCTCTGGGCCACGAGAAACAGTCTCCTCAGCTGCGAGGAGGAGGAGGAGGAGG	959
Db	916	ATAGTCTCTGGGTCCACGAGAAACAGTCTCCTCAGCTGCGAGGAGGAGGAGGAGGAGG	975
QY	960	GATCTGCAGATGGGGCCAAAGATTACTTTCGAAATGAGGAGGAGGAGGAGGAGGAGG	1019
Db	976	GATCTGCAGATGGGGCCAAAGATTACTTTCGAAATGAGGAGGAGGAGGAGGAGGAGG	1035
QY	1020	TGCTGTGCTTAAGGGAAGGGAATCAGAAAGGTGAGGAGGAGGAGGAGGAGGAGGAGG	1079
Db	1036	TGCTGTGCTTAAGGGAAGGGAATCAGAAAGGTGAGGAGGAGGAGGAGGAGGAGGAGG	1095
QY	1080	CCAAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAGTAACTCTCTGAGATTGAAC	1139
Db	1096	CCAAAGAAAGAGTTTTCAGATTTTGAAGAAAGAAACAGTAACTCTCTGAGATTGAAC	1155
QY	1140	AGACAGAGGGGACACAGAGAAAGAGAAATGATCAAGAGAAAGGCGCGAGGAGCTTGA	1199
Db	1156	AGACAGAGGGGACACAGAGAAAGAGAAATGATCAAGAGAAAGGCGCGAGGAGCTTGA	1215
QY	1200	GCGAAGTGGAGACACTGAACTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAG	1259
Db	1216	GCGAAGTGGAGACACTGAACTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAG	1275
QY	1260	ATACAAATCTCAGCGAGGAGCTGAGCTTAATGAGAAAGAGACTTCAAGAAAGTGTCA	1319
Db	1276	ATACAAATCTCAGCGAGGAGCTGAGCTTAATGAGAAAGAGACTTCAAGAAAGTGTCA	1335
QY	1320	TTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1379
Db	1336	TTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1395
QY	1380	CTAACAAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAA	1439
Db	1396	CTAACAAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAA	1455
QY	1440	CTAAACACAGAGGATGAAAGTCCAAATTAATCTGTGTACAGATGACACAAAGGCTT	1499
Db	1456	CTAAACACAGAGGATGAAAGTCCAAATTAATCTGTGTACAGATGACACAAAGGCTT	1515
QY	1500	TTCAAGAACATATAATGCAATTTGAAACCAATTCAGGAGGAGGAGGAGGAGGAGGAGG	1559
Db	1516	TTCAAGAACATATAATGCAATTTGAAACCAATTCAGGAGGAGGAGGAGGAGGAGGAGG	1575
QY	1560	AAGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1619
Db	1576	AAGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1635
QY	1620	TGCTTTCCAAACAGCTGCAAAATGGAATGAAGCAAAACCAATTCGCAAGCAGGAGG	1679
Db	1636	TGCTTTCCAAACAGCTGCAAAATGGAATGAAGCAAAACCAATTCGCAAGCAGGAGG	1695
QY	1680	ACCTGGAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATG	1739
Db	1696	ACCTGGAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATG	1755
QY	1740	CTGAAAGAGCAGCGAGAGAGAGAAATTCATGAGGAAAGGAGGAGGAGGAGGAGGAGG	1799
Db	1756	CTGAAAGAGCAGCGAGAGAGAGAAATTCATGAGGAAAGGAGGAGGAGGAGGAGGAGG	1815
QY	1800	CAGTTCTGCTGAAGAGAGATGATGCTTTTGAAGACGAGGAGGAGGAGGAGGAGGAGG	1859
Db	1816	CAGTTCTGCTGAAGAGAGATGATGCTTTTGAAGACGAGGAGGAGGAGGAGGAGGAGG	1875

QY 1860 TGCAGAGTCGTATGCGGGCGAGAAACAAGTACTCTGACAGCAGCGCTTACCTTGTTCAAA 1919
DB 1876 TGCAGAGTCGTATGCGGGCGAGAAACAAGTACTCTGACAGCAGCGCTTACCTTGTTCAAA 1935
QY 1920 GAGGAGCTGAGACAGGAGCTGGCGGCAACACGCGGAATATCCGATTCATTCCTGCCCA 1979
DB 1936 GAGGAGCTGAGACAGGAGCTGGCGGCAACACGCGGAATATTCGATTCATTCCTGCCCA 1995
QY 1980 AGTGTGGAGAGCTTTCGCTGATAGACATAGACACGCTTACAGATTCAGATTCGATCA 2039
DB 1996 AGTGTGGAGAGCTTTCGCTGATAGACATAGACACGCTTACAGATTCAGATTCGATCA 2055
QY 2040 TTTAAGTGTGTATGATATCACTCCCAAACTGTTGGT 2077
DB 2056 TTTAAGTGTGTATGATATCACTCCCAAACTGTTGGT 2093

RESULT 4
US-10-717-597-79
; Sequence 79, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dörner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79
; LENGTH: 2139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-717-597-79

Query Match 99.0%; Score 2056.4; DB 8; Length 2139;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2071; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1 ATCCCGGTGGAGTTCCTCCAGGCGGCACGATCCGAGGAAACAGTGACCTCGAGCA 60
DB 16 ATCCCGGTGGAGTTCCTCCAGGCGGCACGATCCGAGGAAACAGTGACCTCGAGCA 75
QY 61 AGCCAAAGCGGCGGAGGTGTGGCTTTGATAGCTGGTGGTCCCACTTCCTGGCCCTTGA 120
DB 76 ACCCAAGCGGCGGAGGTGTGGCTTTGATAGCTGGTGGTGGCTTCCTGGCCCTTGA 135
QY 121 TGAGCGGTACGCTCTGTAAACCCCACTTCCTCCTACCTTTGAAACAGCTGCTGGTTCAGC 180
DB 136 TGAGCGGTACGCTCTGTAAACCCCACTTCCTCCTACCTTTGAAACAGCTGCTGGTTCAGC 195
QY 181 ATTAATGAGATTAAGTACAGTACAGCGCTGGTGTGCTGAGTCGACATAGAGATCAA 240
DB 196 ATTAATGAGATTAAGTACAGTACAGCGCTGGTGTGCTGAGTCGACATAGAGATCAA 255
QY 241 AAATGTCCAAATGTAACTGGAGAGAAAGTGGGCACTTTT-GGAGTGACTTTTCCACAG 299
DB 256 AAATGTCCAAATGTAACTGGAGAGAAAGTGGGCACTTTTGGAGTGACTTTTCCACAG 315
QY 300 GAACTTCGCAATGTCCCATCAACCTCTCAGCTGCTCACTGAAAGAGGAGCAGCCCA 359
DB 316 GAACTTCGCAATGTCCCATCAACCTCTCAGCTGCTCACTGAAAGAGGAGCAGCCCA 375

QY 360 GTGAAGCACAGGAATGCACCCACCTGGCCCAACCCAACTGGACACGTTTACCC 419
DB 376 GTGAAGCACAGGAATGCACCCACCTGGCCCAACCCAACTGGACACGTTTACCC 435
QY 420 CGGAGAGCTGCTGCAGCAGATGAAGAGCTCTCTGACCAAGAACCAACAGCTGGAAGAG 479
DB 436 CGGAGAGCTGCTGCAGCAGATGAAGAGCTCTCTGACCAAGAACCAACAGCTGGAAGAG 495
QY 480 CCATGAAGCTTAATTAATCAAGCCATGAAGGGAGATTTGAGAGCTTTCCGCTCGACAG 539
DB 496 CCATGAAGCTTAATTAATCAAGCCATGAAGGGAGATTTGAGAGCTTTCCGCTCGACAG 555
QY 540 AGAAACAGAGGAAGAACGCCAGTCTTTTGTAGATACAGAGCAAGCAAGCAAGAGCGTC 599
DB 556 AGAAACAGAGGAAGAACGCCAGTCTTTTGTAGATACAGAGCAAGCAAGAGAGCGTC 615
QY 600 TAATGGCCTTGAGTCAATGAGAAATGAAGGAAGAGCTTGGAAATCTAAAGGGA 659
DB 616 TAATGGCCTTGAGTCAATGAGAAATGAAGGAAGAGCTTGGAAATCTAAAGGGA 675
QY 660 AATCAGAAAGTTCATCTCAGGACCCCACTGATCACTCCAGGCTTCCAGGCGCGAAGCG 719
DB 676 AATCAGAAAGTTCATCTCAGGACCCCACTGATCACTCCAGGCTTCCAGGCGCGAAGCG 735
QY 720 AGCAGGAAAGGACCCAGCTCAGGACCCAGGTGGTCAAGCTCAAGCAGAGAGGAGGAGCC 779
DB 736 AGCAGGAAAGGACCCAGCTCAGGACCCAGGTGGTCAAGCTCAAGCAGAGAGGAGGAGCC 795
QY 780 TGTGGGCTGCTGTCTGAACTGAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 839
DB 796 TGTGGGCTGCTGTCTGAACTGAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 855
QY 840 CCTTTGTTGAAATTAGGATGCTGAGAGAGAGAGAGGCTCAGTAAAGAAATCAAGC 899
DB 856 CCTTTGTTGAAATTAGGATGCTGAGAGAGAGAGAGGCTCAGTAAAGAAATCAAGC 915
QY 900 ATAGTCTGGGCGCACAGAGACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGAGA 959
DB 916 ATAGTCTGGGCGCACAGAGACAGTCTCCACTGGCAGCGCATTTGTCTCAATATAGGAGA 975
QY 960 GATCTGAGATGGGCGCAAGAAATTAATCTCGAAACATGAGAGGATTAATCTGAGCCAGCTCC 1019
DB 976 GATCTGAGATGGGCGCAAGAAATTAATCTCGAAACATGAGAGGATTAATCTGAGCCAGCTCC 1035
QY 1020 TCCTGCTGAAGGAGGAGATCAGAGGTTGGAGAGCTTCAAGCTTGCACCTCAAGGAGC 1079
DB 1036 TGCTGTGCTTAAGGAGGAGGAGATCAGAGGTTGGAGAGCTTGAAGTTGCACCTCAAGGAG 1095
QY 1080 CCAAGAAAGAGTTCAGATTTTGAAGAAAGAAACAAAGTAATCTGTTCTGAGATTGAACCC 1139
DB 1096 CCAAGAAAGAGTTCAGATTTTGAAGAAAGAAACAAAGTAATCTGTTCTGAGATTGAACCC 1155
QY 1140 AGACAGAGGAGGAGCACAGAGAAAGAGATGATGAAGAGAAAGGCGCGGAGAGCTGTTGGAA 1199
DB 1156 AGACAGAGGAGGAGCACAGAGAAAGAGATGATGAAGAGAAAGGCGCGGAGAGCTGTTGGAA 1215
QY 1200 GCGAGTTCGAGACCTGAACTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTC 1259
DB 1216 GCGAGTTCGAGACCTGAACTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTC 1275
QY 1260 ATACAAAATCTCAGCCAAAGCTGAGCTTAATGAAGAGAGAGCTTCAAGAAAGTGTCAAGGCC 1319
DB 1276 ATACAAAATCTCAGCCAAAGCTGAGCTTAATGAAGAGAGAGCTTCAAGAAAGTGTCAAGGCC 1335
QY 1320 TTGAAGGAAATTTCTCAATTCATCAGAGTTGAAATGAAGAGAGAGCTTGTGTTTATA 1379
DB 1336 TTGAAGGAAATTTCTCAATTTCCATCAGAGTTGAAATGAAGAGAGAGCTTGTGTTTATA 1395
QY 1380 CTACAAAAGCTTAGAGCTTCAAGTGGAAAGCATCTATCAAGAAATCAAAATGGAACAGC 1439
DB 1396 CTACAAAAGCTTAGAGCTTCAAGTGGAAAGCATCTATCAAGAAATCAAAATGGAACAGC 1455

Qy	1440	CTAAACAGACAGGATGAAAGTCCAAATTAACTGTGCTACAGATGACACACAAAGCTTC	1499
Db	1456	CTAAACAGAGGATGAAAGTCCAAATTAACTGTGCTACAGATGACACACAAAGCTTC	1515
Qy	1500	TTCAAGAACATAAATAATGCATTGAAACCAATTGAGGAACATAACAAGAAAAAGAGTCAGAAA	1559
Db	1516	TTCAAGAACATAAATAATGCATTGAAACCAATTGAGGAACATAACAAGAAAAAGAGTCAGAAA	1575
Qy	1560	AAGTGGACAGGGCAGTGTCTGAAGGAACTGAGTGAAGAAAACCTGGAACTGGCAGAGAAGGCTC	1619
Db	1576	AAGTGGACAGGGCAGTGTCTGAAGGAACTGAGTGAAGAAAACCTGGAACTGGCAGAGAAGGCTC	1635
Qy	1620	TGGCTTCCAAAACAGCTGCARAAATGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGG	1679
Db	1636	TGGCTTCCAAAACAGCTGCARAAATGATGAAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGG	1695
Qy	1680	ACCTGGAAACCATGACCACTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATG	1739
Db	1696	ACCTGGAAACCATGACCACTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCATG	1755
Qy	1740	CTGAAGAGCAGCGAGAGAGAAAATTCATGAGGAAAAGGACCACTGGCATTGGCAGCTGG	1799
Db	1756	CTGAAGAGCAGCGAGAGAGAAAATTCATGAGGAAAAGGACCACTGGCATTGGCAGCTGG	1815
Qy	1800	CAGTTCTGCTGAAAGAGAAATGATGCTTTCGAAGACGGAGGCGAGGCAGTCTCTTGATGAGA	1859
Db	1816	CAGTTCTGCTGAAAGAGAAATGATGCTTTCGAAGACGGAGGCGAGGCAGTCTCTTGATGAGA	1875
Qy	1860	TGCAGAGTGTCTATGGGCGAGAAACAAGTGACTCTGAACGACGGCTTACTTGTTCAAA	1919
Db	1876	TGCAGAGTGTCTATGGGCGAGAAACAAGTGACTCTGAACGACGGCTTACTTGTTCAAA	1935
Qy	1920	GAGGAGCTGAGGACAGGGACTGGCGGCAACAGCGGAATATTCGGATTCATTCCTGCCCCA	1979
Db	1936	GAGGAGCTGAGGACAGGGACTGGCGGCAACAGCGGAATATTCGGATTCATTCCTGCCCCA	1995
Qy	1980	AGTGTGGAGAGGTTCTGCTGCACATAGACAGCTTTACAGATTCAAGTTCATGATGATGATCA	2039
Db	1996	AGTGTGGAGAGGTTCTGCTGCACATAGACAGCTTTACAGATTCAAGTTCATGATGATGATCA	2055
Qy	2040	TTTAAAGTGTGATGATGATCACTCTCCCAAAACTGTTGTT	2077
Db	2056	TTTAAAGTGTGATGATGATCACTCTCCCAAAACTGTTGTT	2093

RESULT 5

US-10-756-149-2698
; Sequence 2698, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Azi2, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2698
; LENGTH: 2139
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-2698

Query March 99.0%; Score 2056.4; DB 10; Length 2139;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2071; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 ATCCCGTTCGGAGTTCTCTCAGGCGGCACGATCCCGAGGAAACAGTGCACCTTGAGCGA 60
 Db 16 ATCCCGTTCGGAGTTCTCTCAGGCGGCACGATCCCGAGGAAACAGTGCACCTTGAGCGA 75

Qy	61	AGCCAAGCCGGCGCGCAGGTGTGCTTTTGATAGCTGTGGTGCCACTTCTCGGCGCTTGGAA	120
Db			
Qy	76	AGCCAAGCCGGCGCGCAGGTGTGCTTTGATAGCTGTGGTGCCACTTCTCGGCGCTTGGAA	135
Db			
Qy	121	TGAGCGGTACGCGCTCTGTAAACCCAACTTCTCTCACTTTGAAACAGCTGCGCTGGTTACAG	180
Db			
Qy	136	TGAGCGGTACGCGCTCTGTAAACCCAACTTCTCTCACTTTGAAACAGCTGCGCTGGTTACAG	195
Db			
Qy	181	ATTAATGAAGATTAGTCAGTGACAGGCGCTGTGTGCTGAGTCCGACATAGAAGAATCAA	240
Db			
Qy	196	ATTAATGAAGATTAGTCAGTGACAGGCGCTGTGTGCTGAGTCCGACATAGAAGAATCAA	255
Db			
Qy	241	AAATGTCCAAAATGTAACTGAGAGAAATGCGGCAACTTTTGGAGTACTTTTCCACAG	299
Db			
Qy	256	AAATGTCCAAAATGTAACTGAGAGAAATGCGGCAACTTTTGGAGTACTTTTCCACAG	315
Db			
Qy	300	GAACTTCTGCAATGTCCTCAACCTCTCAGCTGCTCCTCACTGAAAGAGGACGACGCCCA	359
Db			
Qy	316	GAACTTCTGCAATGTCCTCAACCTCTCAGCTGCTCCTCACTGAAAGAGGACGACGCCCA	375
Db			
Qy	360	GTGAAAGCACAGGAAATGGACCCGCCCACTTGCGGCCACCCAAACCTGGACACGTTTACCC	419
Db			
Qy	376	GTGAAAGCACAGGAAATGGACCCGCCCACTTGCGGCCACCCAAACCTGGACACGTTTACCC	435
Db			
Qy	420	CGGAGGAGTCTCGACGAGATGAAAGAGCTCTTGACCAAGAAACCAACGCTGAAAGAG	479
Db			
Qy	436	CGGAGGAGTCTCGACGAGATGAAAGAGCTCTTGACCAAGAAACCAACGCTGAAAGAG	495
Db			
Qy	480	CCATGAAGCTAAATAATCAAGCCATGAAAGGAGATTTGAGGAGCTTTTGGCGCTTGACAG	539
Db			
Qy	496	CCATGAAGCTAAATAATCAAGCCATGAAAGGAGATTTGAGGAGCTTTTGGCGCTTGACAG	555
Db			
Qy	540	AGAAACAGAGGAAAGACCGCCAGTTTTTCAGATACAGAGCAAAAGAACAAAGAGCGTC	599
Db			
Qy	556	AGAAACAGAGGAAAGACCGCCAGTTTTTCAGATACAGAGCAAAAGAACAAAGAGCGTC	615
Db			
Qy	600	TAATGGCGCTTGAGTCATGAGAAATGAGAAATTTGAAGGAAGAGCTTGGAAAACTAAAGGGA	659
Db			
Qy	616	TAATGGCGCTTGAGTCATGAGAAATGAGAAATTTGAAGGAAGAGCTTGGAAAACTAAAGGGA	675
Db			
Qy	660	AATCAGAAAGGTCACTGAGGACCCCACTGATGACTCCAGGCTTTCCAGGCGCGAGCGG	719
Db			
Qy	676	AATCAGAAAGGTCACTGAGGACCCCACTGATGACTCCAGGCTTTCCAGGCGCGAGCGG	735
Db			
Qy	720	AGCAGGAAAGGACCCAGCTCAGGACCCAGGTGTGAGGCTACAAGCAGAGAAGGCAGACC	779
Db			
Qy	736	AGCAGGAAAGGACCCAGCTCAGGACCCAGGTGTGAGGCTACAAGCAGAGAAGGCAGACC	795
Db			
Qy	780	TGTTGGGCATCGTGTCTGAACTGCGACTCAAGCTGAACTCAGCGGCTCTCAGAGAATT	839
Db			
Qy	796	TGTTGGGCATCGTGTCTGAACTGCGACTCAAGCTGAACTCAGCGGCTCTCAGAGAATT	855
Db			
Qy	840	CTTTTGTGAAATTTAGGATGGCTGAAAGGAGAGCAGAGGCTCAGTAAAGAAATCAAGC	899
Db			
Qy	856	CTTTTGTGAAATTTAGGATGGCTGAAAGGAGAGCAGAGGCTCAGTAAAGAAATCAAGC	915
Db			
Qy	900	ATAGTCTCGGCCCAACAGAACAGCTCTCACTGGCAAGGCTTCTTAAATATAGGAGCA	959
Db			
Qy	916	ATAGTCTCGGTCACAGAGAACAGCTCTCACTGGCACGGCATTTGTCTCACTATAGGAGGA	975
Db			
Qy	960	GATCTGCAGATGGGGCCCAAGAAATTAATTCGAACATGAGGAGTTAACTGTGAGCCAGCTCC	1019
Db			
Qy	976	GATCTGCAGATGGGGCCCAAGAAATTAATTCGAACATGAGGAGTTAACTGTGAGCCAGCTCC	1035
Db			
Qy	1020	TGCTGTGCTTAAGGGAAGGAAATCAGAAAGGTGGAGAGACTTGAAGTTGCACCTCAAGGAGG	1079
Db			
Qy	1036	TGCTGTGCTTAAGGGAAGGAAATCAGAAAGGTGGAGAGACTTGAAGTTGCACCTCAAGGAGG	1095
Db			
Qy	1080	CCAAAGAAAGAGTTTTCAGATTTTTCGAAAGAAAAACAAGTAAATCGTTCTGAGATTGAAACCC	1139
Db			
Qy	1096	CCAAAGAAAGAGTTTTCAGATTTTTCGAAAGAAAAACAAGTAAATCGTTCTGAGATTGAAACCC	1155
Db			
Qy	1140	AGACAGAGGGGAGCACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGAGCTGTGTGAA	1199

Db 1156 ACACAGAGGGAGCAGACAGAGAAAGAGATGATGAAGAGAAAGCGCCGGAGACTGTTGGAA 1215
QY 1200 GCGAAGTGAAGACACTGAACCTCCAGGTGCACATCTCTGTTTAAAGAGCTTCAAGAGGCTC 1259
Db 1216 GCGAAGTGAAGACACTGAACCTCCAGGTGCACATCTCTGTTTAAAGAGCTTCAAGAGGCTC 1275
QY 1260 ATACAAAACCTCAGCAGAGCTGAGCTTAATGAAGAAGAGACTTCAAGAAAAGTGTCAAGGCC 1319
Db 1276 ATACAAAACCTCAGCAGAGCTGAGCTTAATGAAGAAGAGACTTCAAGAAAAGTGTCAAGGCC 1335
QY 1320 TTGAAGGAAAAATTTCTGCAATTTCCATCAGAGTTGAATGAAAAGCAAGAGCTTGTTTATA 1379
Db 1336 TTGAAGGAAAAATTTCTGCAATTTCCATCAGAGTTGAATGAAAAGCAAGAGCTTGTTTATC 1395
QY 1380 CTAACAAAAGTTAGAGCTACAGTGAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGG 1439
Db 1396 CTAACAAAAGTTAGAGCTACAGTGAAGTGAAGAGCATGCTATCAGAAATCAAAATGGAACAGG 1455
QY 1440 CTAACACAGAGATGAAGAGTCCAAATTAACCTGCTACAGATGACACACAACAGCTTC 1499
Db 1456 CTAACACAGAGATGAAGAGTCCAAATTAACCTGCTACAGATGACACACAACAGCTTC 1515
QY 1500 TTCAAGAACATAATATGCAATTTGAAAACAATTTGAGGAACCTAAAGAAAAGAGTCAGAAA 1559
Db 1516 TTCAAGAACATAATATGCAATTTGAAAACAATTTGAGGAACCTAAAGAAAAGAGTCAGAAA 1575
QY 1560 AAGTGAAGAGGAGTGTGTAAGGAACCTGAGTGAAGAACTGAACTGGAACCTGGCAGAGAGGCTC 1619
Db 1576 AAGTGAAGAGGAGTGTGTAAGGAACCTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTC 1635
QY 1620 TGCGTTTCAAAACAGCTGCAAAATGGAATGAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGG 1679
Db 1636 TGCGTTTCAAAACAGCTGCAAAATGGAATGAATGAAGCAAAACCATTTGCCAAGCAGGAAGAGG 1695
QY 1680 ACCTGGAACCCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCAATG 1739
Db 1696 ACCTGGAACCCATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTCAATG 1755
QY 1740 CTGAAGAGCAGGAGAGAGAAAATTCATGAGGAAAAGAGCAACTGGCATTTGCAGCTGG 1799
Db 1756 CTGAAGAGCAGGAGAGAGAAAATTCATGAGGAAAAGAGCAACTGGCATTTGCAGCTGG 1815
QY 1800 CAGTTCCTGTGAAGAGAGATGATCTTTCGAAGCAGGAGGAGGAGTCTTGTGATGGAGA 1859
Db 1816 CAGTTCCTGTGAAGAGAGATGATCTTTCGAAGCAGGAGGAGGAGTCTTGTGATGGAGA 1875
QY 1860 TGCAGAGTCTCATGGGGCGAGAACAAAGTGACTCTGACCCAGCAGGCTTACCTTGTTCAAA 1919
Db 1876 TGCAGAGTCTCATGGGGCGAGAACAAAGTGACTCTGACCCAGCAGGCTTACCTTGTTCAAA 1935
QY 1920 GAGGAGCTGAGNACAGGAGCTGGCGGCAACAGCGGAATATTCGGATTCAATTCCTGCCCA 1979
Db 1936 GAGGAGCTGAGNACAGGAGCTGGCGGCAACAGCGGAATATTCGGATTCAATTCCTGCCCA 1995
QY 1980 AGTGTGAGAGGTTCTGCTGACATAGACAGCTTACAGATTACAGTATGAGGATTCGATCA 2039
Db 1996 AGTGTGAGAGGTTCTGCTGACATAGACAGCTTACAGATTACAGTATGAGGATTCGATCA 2055
QY 2040 TTTAAGTGTGTATGATACCTCCCAAAACTGTTGGT 2077
Db 2056 TTTAAGTGTGTATGATACCTCCCAAAACTGTTGGT 2093

RESULT 6
US-10-609-133-5
; Sequence 5, Application US/10609133
; Publication No. US20040191798A1
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut Health Center
; APPLICANT: Sarfarazi, Mansoor
; APPLICANT: Rezaie, Tayebbeh
; APPLICANT: Child, Anne H

; TITLE OF INVENTION: Optineurin and Glaucoma
; FILE REFERENCE: UCT-0046P2
; CURRENT APPLICATION NUMBER: US/10/609,133
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: 10,281,457
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 10,090,118
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 10,060,981
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: 60/344,754
; PRIOR FILING DATE: 2001-12-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; AUTHORS: Rezaie et al.
; TITLE: Adult-Onset Primary Open-Angle Glaucoma Caused by Mutations in
; TITLE: Optineurin
; JOURNAL: Science
; VOLUME: 295
; ISSUE: 5557
; PAGES: 1077-1079
; DATE: 2002
; DATABASE ACCESSION NUMBER: AF420373
; DATABASE ENTRY DATE: 2002-02-11
; RELEVANT RESIDUES: (1)..(2008)
US-10-609-133-5
Query Match 92.8%; Score 1927.4; DB 9; Length 2008;
Best Local Similarity 96.6%; Pred. No. 0; Mismatches 1; Indels 69; Gaps 1;
Matches 2007; Conservative 0;
QY 1 ATCCCGGTGGGAGTTCTCTCCAGCGCGCAGCATCCGAGGAAACAGTGACCTGAGCGA 60
Db 1 ATCCCGGTGGGAGTTCTCTCCAGCGCGCAGCATCCGAGGAAACAGTGACCTGAGCGA 60
QY 61 AGCCAAAGCCGGCGGCGAGGTGGCTTTGATAGCTGGTGGTCCACTTCTGGCCTTGA 120
Db 61 AGCCAAAGCCGGCGGCGAGGTGGCTTTGATAGCTGGTGGTCCACTTCTGGCCTTGA 120
QY 121 TGAGCCGTACGCTCTGTAAACCCAACTTCCTCACCTTTTGAACAGCTGCTGGTTCAGC 180
Db 121 TGAGCCGTACGCTCTGTAAACCCAACTTCCTCACCTTTTGAACAGCTGCTGGTTCAGC 180
QY 181 ATTAATGAAGATTAGTCAGTGACAGCCCTGGTGTGCTGAGTCCGACATAGAAGATCAA 240
Db 181 ATTAATGAAGATTAGTCAGTGACAGCCCTGGTGTGCTGAGTCCGACATAGAAGATCAA 230
QY 241 AAATGTCCAAATGTAACTGGAGAGAAAGTGGGCAACTTTTGGAGTGACTTTTCCACAGG 300
Db 231 -----G 231
QY 301 RACTTCTCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAGGAGGACAGCCCCAG 360
Db 232 AACTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAAGGAGGACAGCCCCAG 291
QY 361 TGAAGACACAGGAAATGGAACCCCACTGGCCCAACCAAACTTGGACACGCTTTACCCC 420
Db 292 TGAAGACACAGGAAATGGAACCCCACTGGCCCAACCAAACTTGGACACGCTTTACCCC 351
QY 421 GGAGAGCTGTGACAGCAGATGAAGAGCTCTCTGACCAAGAACACCACAGCTGAAGAACG 480
Db 352 GGAGAGCTGTGACAGCAGATGAAGAGCTCTCTGACCAAGAACACCACAGCTGAAGAACG 411
QY 481 CATGAAGCTAAATAATCAAGCCATCAAGGGAGATTGAGGAGCTTTCCGCTTGGACAGA 540
Db 412 CATGAAGCTAAATAATCAAGCCATCAAGGGAGATTGAGGAGCTTTCCGCTTGGACAGA 471
QY 541 GAAACAGAGGAAGAACGCCAGTCTTTTGGAGATACAGACCAAGAAAGAACAGAGCGTCT 600

	Best Local Similarity	99.9%;	Pred. No. 0;	Mismatches	1;	Indels	0;	Gaps	0;
	Matches	1780;	Conservative	0;					
Qy	297	CAGGAACTTTGTCAATGTCCCATCAACCTCTCAGCTGCCTCACCTGAAAAGGAGCAGACC	356						
Dd	76	CAGGAACTTCTGCAATGTCCCATCAA	135						
Qy	357	CCAGTGAAGCACAAGGAAATGGACCCCACCCTTGCGGCCACCCAAAACCTTGGACACGTTTA	416						
Dd	136	CCAAGTGAAGCACAAGGAAATGGACCCCACCCTTGCGGCCACCCAAAACCTTGGACACGTTTA	195						
Qy	417	CCCCGAGGAGCTGCTCAGCAGATGAAGAGCTCCTGACCAAGAACCCACAGCTGAAAG	476						
Dd	196	CCCCGAGGAGCTGCTGACGAGATGAAGAGCTTCCTGACCGAGAACCAACAGCTGAAAG	255						
Qy	477	AAGCCATTGAAGCTTAAATAAACAAGCCCATGAAGGGAGATTTGAGGAGCTTTCGGCCTTGA	536						
Dd	256	AAGCCATTGAAGCTTAAATAAACAAGCCCATGAAGGGAGATTTGAGGAGCTTTCGGCCTTGA	315						
Qy	537	CAGGAAACAGAAGGAAGACCCACAGTTTTTTTGATATACAGAGCAAGAAAGCAAAAGAC	596						
Dd	316	CAGGAAACAGAAGGAAGACCCACAGTTTTTTTGATATACAGAGCAAGAAAGCAAAAGAC	375						
Qy	597	GTCTAAATGGCCTTCAGTTCATGAGAAATGAGAAATGAAGGAAGAGCTTGGAAAACTAAAG	656						
Dd	376	GTCTAAATGGCCTTCAGTTCATGAGAAATGAGAAATGAAGGAAGAGCTTGGAAAACTAAAG	435						
Qy	657	GGAAATCAGAAAGTCTATCTGAGAACCCCACTGATGATCTCCAGGCTTCCACAGGCGCAAG	716						
Dd	436	GGAAATCAGAAAGTCTATCTGAGAACCCCACTGATGATCTCCAGGCTTCCACAGGCGCAAG	495						
Qy	717	CGGAGCAGGAAAAGGACACAGCTCAGGACCCACAGTGTGAGCTACAAAGCAGAGAAGGCAG	776						
Dd	496	CGGAGCAGGAAAAGGACACAGCTCAGGACCCACAGTGTGAGCTACAAAGCAGAGAAGGCAG	555						
Qy	777	ACTGTTTGGGCATCGTGTCTGAACTGACGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAG	836						
Dd	556	ACTGTTTGGGCATCGTGTCTGAACTGACGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAG	615						
Qy	837	ATTCCTTTGTGAAATTAGNTGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA	896						
Dd	616	ATTCCTTTGTGAAATTAGNTGCTGAAGGAGAGCAGAGGCTCAGTAAAGAAATCA	675						
Qy	897	AGCATAGTCTCTGGGCCACAGAAACAGTCTCCACTGGCACGGCATTTCTAAATATAGGA	956						
Dd	676	AGCATAGTCTCTGGGCCACAGAAACAGTCTCCACTGGCACGGCATTTCTAAATATAGGA	735						
Qy	957	GCAGATCTGCAGATGGGGCCAGAAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGC	1016						
Dd	736	GCAGATCTGCAGATGGGGCCAGAAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGC	795						
Qy	1017	TCTGCTGTGCTTAAGGGMAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG	1076						
Dd	796	TCTGCTGTGCTTAAGGGMAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG	855						
Qy	1077	AGGCCAAAGAAAGATTTTCAGATTTTGAAGAAAGAAACAAGTAATCGTTCTTGAGATTGAAA	1136						
Dd	856	AGGCCAAAGAAAGATTTTCAGATTTTGAAGAAAGAAACAAGTAATCGTTCTTGAGATTGAAA	915						
Qy	1137	CCAGACAGAGGGGAGCACAGNAAAGAGATGATGAAGAAAGGCCCGGAGACTCTTG	1196						
Dd	916	CCAGACAGAGGGGAGCACAGNAAAGAGATGATGAAGAAAGGCCCGGAGACTCTTG	975						
Qy	1197	GAAGCGAAGTGAAGAGCACTGAACCTCCAGGCTGACATCTGTGTTTAAGGAGCTTCAAGAGG	1256						
Dd	976	GAAGCGAAGTGAAGAGCACTGAACCTCCAGGCTGACATCTGTGTTTAAGGAGCTTCAAGAGG	1035						
Qy	1257	CTCATACAAACTCAGCGAGCTAGCTAATGAAGAGAGACTTTCAAGAAAGTGTCTAGG	1316						
Dd	1036	CTCATACAAACTCAGCGAGCTAGCTAATGAAGAGAGACTTTCAAGAAAGTGTCTAGG	1095						
Qy	1317	CCCTTTGAAGGAAAAATTTCTCAATTCATCAGAGTTGAATGAAAAAGCAAGAGCTTCTGTT	1376						

1096	CCCTTGAAGGAAAAATTCTTGCAAATTCACATCAGAGTTGAAATGAAAAACCAAGAGCGTTGTTTT	1155
Db		
1377	ATACTAACAAAAAGTTAGAGCTCAAGATGTGAAAGCATGCTTATCAGAAATCAAAATGGAAC	1436
Qy		
Db		
1156	ATACTACAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC	1215
Qy		
1437	AGGCTAAAAACAGAGGATGAAAAGTCCAAAATTAACTGTGTCTACAGATGACACAAACAAGC	1496
Db		
1216	AGGCTAAAAACAGAGGATGAAAAGTCCAAAATTAACTGTGTCTACAGATGACACAAACAAGC	1275
Qy		
1497	TTCTTTCAGAACATATAATAGTCATTGAAACCAATTGAGGAACCTAAACAAGAAAAGAGTTCAG	1556
Db		
1276	TTCTTTCAGAACATATAATAGTCATTGAAACCAATTGAGGAACCTAAACAAGAAAAGAGTTCAG	1335
Qy		
1557	AAAAAGTGGACAGGGCAGTCTGAAGAACTGAGTGA AAAAATCTGGAATCTGGCAGAGAAGG	1616
Db		
1336	AAAAAGTGGACAGGGCAGTCTGAAGAACTGAGTGA AAAAATCTGGAATCTGGCAGAGAAGG	1395
Qy		
1617	CTCTGGCTTCCAAACAGCTCAAAATGGATGAAATGAAGCAAAACCAATGCCCAGCAGGAAG	1676
Db		
1396	CTCTGGCTTCCAAACAGCTCAAAATGGATGAAATGAAGCAAAACCAATGCCCAGCAGGAAG	1455
Qy		
1677	AGNACTGGAAAACCATGACCATCTCTAGGCGCTCAGATGGAAGTTTACTGTTCGATTTTC	1736
Db		
1456	AGGACCTGGAAAACCATGACCATCTCTAGGCGCTCAGATGGAAGTTTACTGTTCGATTTTC	1515
Qy		
1737	ATGCTGAAGAGCAGCGAGAGAGAAAAATTCAATGAGGAAAAAGGAGCAACTGGCATTTGCAGC	1796
Db		
1516	ATGCTGAAGAGCAGCGAGAGAGAAAAATTCAATGAGGAAAAAGGAGCAACTGGCATTTGCAGC	1575
Qy		
1797	TGGCAGTTCTGCTGAAAGAGAATGATGCTTTTCAAGACGGAGGCAGCAGTCTCTTGATGG	1856
Db		
1576	TGGCAGTTCTGCTGAAAGAGAATGATGCTTTTCAAGACGGAGGCAGCAGTCTCTTGATGG	1635
Qy		
1857	AGATGCAGAGTCTGATGGGGCAGAAACAAGTGACTCTGACCCAGCAGGCTTACCTTGTTTC	1916
Db		
1636	AGATGCAGAGTCTGATGGGGCAGAAACAAGTGACTCTGACCCAGCAGGCTTACCTTGTTTC	1695
Qy		
1917	AAAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAAATATTCCGATTCATTTCCTGCC	1976
Db		
1696	AAAGAGAGCTGAGGACAGGAGCTGGCGGCAACAGCGGAAATATTCCGATTCATTTCCTGCC	1755
Qy		
1977	CCAAGTGTGGAGAGTTCTCGCTGCACATAGACAGTTTACAGATTCACGTGATGGATTGCA	2036
Db		
1756	CCAAGTGTGGAGAGTTCTCGCTGCACATAGACAGTTTACAGATTCACGTGATGGATTGCA	1815
Qy		
2037	TCATTTAAAGTGTGATGATATCACCTCTCCCAAAACCTGTTGGT	2077
Db		
1816	TCATTTAAAGTGTGATGATATCACCTCTCCCAAAACCTGTTGGT	1856

```

RESULT 8
US-10-755-889-213
; Sequence 213, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 213
; LENGTH: 2327
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-213

```


Query Match		85.7%;	Score 1779.4;	DB 8;	Length 2327;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 1780;		Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
Qy	297	CAGGAACTTCTGCAATCTCCATCAACCTCTCAGCTGCCTCAGTGAAGAGGAGACGCC	356			
Db	141	CAGGAACTTCTGCAATCTCCATCAACCTCTCAGCTGCCTCAGTGAAGAGGAGACGCC	200			
Qy	357	CCAGTGAAGCACAGGAAATGGACCCGCCCACTGGGCCACCCCAAACTCTGGACACGTTTA	416			
Db	201	CCAGTGAAGCACAGGAAATGGACCCGCCCACTGGGCCACCCCAAACTCTGGACACGTTTA	260			
Qy	417	CCCAGGAGGCTGCTCAGCAGATGAAGAGCTCCTGACCAAGAACCCACAGCTGAAG	476			
Db	261	CCCAGGAGGCTGCTCAGCAGATGAAGAGCTCCTGACCGAGAACCCACAGCTGAAG	320			
Qy	477	AAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTCGGCCTGGA	536			
Db	321	AAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTCGGCCTGGA	380			
Qy	537	CAGAGAAACAGAGGAAGACGCCAGCTTTTGTGATACAGAGCAAGAGAGCAAAAGAGC	596			
Db	381	CAGAGAAACAGAGGAAGACGCCAGCTTTTGTGATACAGAGCAAGAGAGCAAAAGAGC	440			
Qy	597	GTCTAATGGCTTGAGTCATGAGAAATGAGAAATTTGAAGGAAGAGCTTTGGAATACTAAAG	656			
Db	441	GTCTAATGGCTTGAGTCATGAGAAATGAGAAATTTGAAGGAAGAGCTTTGGAATACTAAAG	500			
Qy	657	GGAAATCAGAAAGGTCTCTGAGGACCCCACTGATGATCTCAGGCTTCCAGGGCCGAAG	716			
Db	501	GGAAATCAGAAAGGTCTCTGAGGACCCCACTGATGATCTCAGGCTTCCAGGGCCGAAG	560			
Qy	717	CGAGCAGGAAAGGACACAGCTCAGGACCCAGGTGGTGGCTACAGCAGAGAGAGGACG	776			
Db	561	CGAGCAGGAAAGGACACAGCTCAGGACCCAGGTGGTGGCTACAGCAGAGAGAGGACG	620			
Qy	777	ACCTGTTGGGCTCGTGTCTGAACTGAGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAG	836			
Db	621	ACCTGTTGGGCTCGTGTCTGAACTGAGCTCAAGCTGAACTCCAGCGGCTCCTCAGAAG	680			
Qy	837	ATTCTTTGTTGAAATTTAGATGGCTGAAGAGAGACAGAGGTCAGTAAAGAAATCA	896			
Db	681	ATTCTTTGTTGAAATTTAGATGGCTGAAGAGAGACAGAGGTCAGTAAAGAAATCA	740			
Qy	897	AGCATAGTCTGGGCCCAACAGAACAGTCTCCACTGGCAGCGATTCTCTAAATATAGGA	956			
Db	741	AGCATAGTCTGGGCCCAACAGAACAGTCTCCACTGGCAGCGATTCTCTAAATATAGGA	800			
Qy	957	GCAGATCTGCAGATGGGCCCAAGAAATTAATTCGAACATAGAGGATTAACCTGTGAGCCAGC	1016			
Db	801	GCAGATCTGCAGATGGGCCCAAGAAATTAATTCGAACATAGAGGATTAACCTGTGAGCCAGC	860			
Qy	1017	TCCTGCTGCTTAAGGGAAAGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG	1076			
Db	861	TCCTGCTGCTTAAGGGAAAGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG	920			
Qy	1077	AGGCCAAGAAAGAGTTTCAGATTTTGAABAGAAACAAGTAATCGTTCTGAGATTGAAA	1136			
Db	921	AGGCCAAGAAAGAGTTTCAGATTTTGAABAGAAACAAGTAATCGTTCTGAGATTGAAA	980			
Qy	1137	CCCAGACAGAGGGGAGCACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTG	1196			
Db	981	CCCAGACAGAGGGGAGCACAGAGAAAGAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTG	1040			
Qy	1197	GAGCGAAGTGGAGACCTGAACCTCCAGTGACATCTCTGTTTAAGGAGCTTCAAGAGG	1256			
Db	1041	GAGCGAAGTGGAGACCTGAACCTCCAGTGACATCTCTGTTTAAGGAGCTTCAAGAGG	1100			
Qy	1257	CTCATACAAACTCAGCGAGCTGAGCTAATGAGAGAGACTTCAGAAAGAGTGCAGG	1316			
Db	1101	CTCATACAAACTCAGCGAGCTGAGCTAATGAGAGAGACTTCAGAAAGAGTGCAGG	1160			

Qy	1317	CCCTTGAAGGAAAAATTTCTGCAATTCATCCATCAGAGTTGAATGAAAAAGCAAGAGCTTGTTT	1376			
Db	1161	CCCTTGAAGGAAAAATTTCTGCAATTCATCCATCAGAGTTGAATGAAAAAGCAAGAGCTTGTTT	1220			
Qy	1377	ATACTAACAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC	1436			
Db	1221	ATACTAACAAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC	1280			
Qy	1437	AGGCTAAAAACAGAGGATGAAAAGTCCAAATTTAACTGTCTACAGATTGACACAAACAAGC	1496			
Db	1281	AGGCTAAAAACAGAGGATGAAAAGTCCAAATTTAACTGTCTACAGATTGACACAAACAAGC	1340			
Qy	1497	TTCTTCAAGAACATAAATGCAATGAAAACAAATGAGGAATTAACAAGAAAAAGAGTCAG	1556			
Db	1341	TTCTTCAAGAACATAAATGCAATGAAAACAAATGAGGAATTAACAAGAAAAAGAGTCAG	1400			
Qy	1557	AAAAAGTGCACAGGGCAGTGTGAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGG	1616			
Db	1401	AAAAAGTGCACAGGGCAGTGTGAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGG	1460			
Qy	1617	CTCTGGCTTCCAAACAGCTGCAAAATGGAATGAAGCAAAACCATTTGCCAACGACGAAG	1676			
Db	1461	CTCTGGCTTCCAAACAGCTGCAAAATGGAATGAAGCAAAACCATTTGCCAACGACGAAG	1520			
Qy	1677	AGGACCTGSAACCATATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTC	1736			
Db	1521	AGGACCTGSAACCATATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTCTGATTTTC	1580			
Qy	1737	ATGCTGAAGAGCAGCGAGAGAGAAAATTCATGAGAAAGGAGCAACTGGCAATTCGAGC	1796			
Db	1581	ATGCTGAAGAGCAGCGAGAGAGAAAATTCATGAGAAAGGAGCAACTGGCAATTCGAGC	1640			
Qy	1797	TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTCGAAGACGGAGCAGGCACTCCTTGATGG	1856			
Db	1641	TGGCAGTTCTGCTGAAAGAGAAATGATGCTTTCGAAGACGGAGCAGGCACTCCTTGATGG	1700			
Qy	1857	AGATGACAGTCTGATGGGGGAGAAACAAAGTGAATCTCTGACAGAGGCTTACCTTGTTTC	1916			
Db	1701	AGATGACAGTCTGATGGGGGAGAAACAAAGTGAATCTCTGACAGAGGCTTACCTTGTTTC	1760			
Qy	1917	AAAGAGGAGCTGAGGACAGGAGCTGGGGGCAACAGCGGAATATTCGATTCATTCCTGCC	1976			
Db	1761	AAAGAGGAGCTGAGGACAGGAGCTGGGGGCAACAGCGGAATATTCGATTCATTCCTGCC	1820			
Qy	1977	CCAAGTGTGGAGAGTTCTGCTGACATAGACAGTTACAGATTACAGTGTGATGCA	2036			
Db	1821	CCAAGTGTGGAGAGTTCTGCTGACATAGACAGTTACAGATTACAGTGTGATGCA	1880			
Qy	2037	TCATTTAAGTGTGATGATACCTCCCAAACTGTTGGT	2077			
Db	1881	TCATTTAAGTGTGATGATACCTCCCAAACTGTTGGT	1921			

RESULT 9

US-10-956-157-2378
; Sequence 2378, Application US/10956157
; Publication No. US20050118625A1

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Mounts, William

; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH

; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES

; FILE REFERENCE: 031896-043000 (AM 101081)

; CURRENT APPLICATION NUMBER: US/10/956,157

; CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 319805

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 2378

; LENGTH: 2327

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-956-157-2378

Query Match				85.7%;	Score 1779.4;	DB 10;	Length 2327;				
Best Local Similarity				99.9%;	Pred. No. 0;						
Matches 1780;				Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Q	297	CAGGAACCTTCTGCAATGTCCTCAATCAACCTCTCAGCTGCTCTCACTGAAAGAGGACAGCC	356								
Q	141	CAGGAACCTTCTGCAATGTCCTCAATCAACCTCTCAGCTGCTCTCACTGAAAGAGGACAGCC	200								
Q	357	CCAGTGAAGACACAGGAATGGACCCGCCACCTGGCCACCCCAAACTCGACACGTTT	416								
Q	201	CCAGTGAAGACACAGGAATTGAACCCGCCACCTGGCCACCCCAAACTCGACACGTTT	260								
Q	417	CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCCTGACCAAGAACCAACGCTGAAG	476								
Q	261	CCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCCTGACCCGAGAACCAACGCTGAAG	320								
Q	477	AAGCCATGAAGCTAAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTCCGCTGGA	536								
Q	321	AAGCCATGAAGCTAAATATCAAGCCATGAAGGGAGATTTGAGGAGCTTTCCGCTGGA	380								
Q	537	CAGAGAAACAGAGGAAGAACCGCAGATTTTGTAGATACAGAGCAAAAGAACGAGGAGC	596								
Q	381	CAGAGAAACAGAGGAAGAACCGCAGATTTTGTAGATACAGAGCAAAAGAACGAGGAGC	440								
Q	597	GTCTAATGGCCTTGAGTCAATGAGAAATGAAGAAATTTGAAGGAGAGCTTTGAAAACCTAAAG	656								
Q	441	GTCTAATGGCCTTGAGTCAATGAGAAATGAAGAAATTTGAAGGAGAGCTTTGAAAACCTAAAG	500								
Q	657	GGAAATCAGAAAGGTCAATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAG	716								
Q	501	GGAAATCAGAAAGGTCAATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAG	560								
Q	717	CGGAGCAGAAAGGACACAGCTCAGGACCCAGGTGCTCAGGCTACAGCAGAGAGGCGAG	776								
Q	561	CGGAGCAGAAAGGACACAGCTCAGGACCCAGGTGCTCAGGCTACAGCAGAGAGGCGAG	620								
Q	777	ACCTGTTGGGCATCGTGTCTGAACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAAG	836								
Q	621	ACCTGTTGGGCATCGTGTCTGAACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAAG	680								
Q	837	ATTCCTTTTGAATTAAGATGCTGAAGGAGAGCAGAAAGGTCAGTAAAGAAATCA	896								
Q	681	ATTCCTTTTGAATTAAGATGCTGAAGGAGAGCAGAAAGGTCAGTAAAGAAATCA	740								
Q	897	AGCATAGTCTGGGCCACAGAGACAGTCTCCACTGGCAGCGCATGCTCTAATATAGA	956								
Q	741	AGCATAGTCTGGGCCACAGAGACAGTCTCCACTGGCAGCGCATGCTCTAATATAGA	800								
Q	957	GCAGATCTGCAGATGGGGCCAAAGAAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGC	1016								
Q	801	GCAGATCTGCAGATGGGGCCAAAGAAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGC	860								
Q	1017	TCCTGCTGTGCTTAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG	1076								
Q	861	TCCTGCTGTGCTTAAGGGAAGGGAATCAGAAAGTGGAGAGACTTGAAGTTGCACTCAAGG	920								
Q	1077	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAACTGTTCTGAGATTTGAAA	1136								
Q	921	AGGCCAAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAACTGTTCTGAGATTTGAAA	980								
Q	1137	CCCAGACAGAGGGGACACAGAGAAAGAGAATGAAGAGAAAGGCCCGGAGACTGTTG	1196								
Q	981	CCCAGACAGAGGGGACACAGAGAAAGAGAATGAAGAGAAAGGCCCGGAGACTGTTG	1040								
Q	1197	GAAGCGAAGTGAAGCACTGAACCTCAGGTCGACATCTCTGTTTAAAGGAGCTTCAAGAGG	1256								
Q	1041	GAAGCGAAGTGAAGCACTGAACCTCAGGTCGACATCTCTGTTTAAAGGAGCTTCAAGAGG	1100								
Q	1257	CTCATACAAACTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAGTGTCAAG	1316								
Q	1101	CTCATACAAACTCAGCGAAGCTGAGCTTAATGAAGAGAGACTTCAAGAAAAGTGTCAAG	1160								
Q	1317	CCCTTGAAAGGAAAAATCTGCAATTCATCAGAGTTGAAATGAAAGCAAGGCTTGTT	1376								

DB	1161	CCCTTGAAAGGAAAAATCTGCATTCATCAGAGTTCAATGAAAGCAAGAGCTTGT	1220								
QY	1377	ATACTAAACAAAAGTTAGAGCTACAGTGGAAAGCATCTATCAGAAATCAAAATGGAAAC	1436								
DB	1221	ATACTAAACAAAAGTTAGAGCTACAGTGGAAAGCATCTATCAGAAATCAAAATGGAAAC	1280								
QY	1437	AGGCTTAAACACAGAGGATGAAAAGTCCAAATTAACCTGTCTACAGATGACACAACAAGC	1496								
DB	1281	AGGCTTAAACACAGAGGATGAAAAGTCCAAATTAACCTGTCTACAGATGACACAACAAGC	1340								
QY	1497	TTCTTCAAGAACATAATATCATTTGAAAAACAATTTAGAGAACTAAACAGAAAAGAGTCAG	1556								
DB	1341	TTCTTCAAGAACATAATATCATTTGAAAAACAATTTAGAGAACTAAACAGAAAAGAGTCAG	1400								
QY	1557	AAAAAGTGAACAGGGCAGTGTGAAGGAACCTGAGTGAAGAACTGGAACCTGGCAGAGAGG	1616								
DB	1401	AAAAAGTGAACAGGGCAGTGTGAAGGAACCTGAGTGAAGAACTGGAACCTGGCAGAGAGG	1460								
QY	1617	CTCTGGCTTCCAAAACAGCTGCAAAATGGAATGAAATGAAGCAAAACCAATTCGAAGCAGGAAG	1676								
DB	1461	CTCTGGCTTCCAAAACAGCTGCAAAATGGAATGGAATGAAGCAAAACCAATTCGAAGCAGGAAG	1520								
QY	1677	AGGACCTGGAACCAATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTC	1736								
DB	1521	AGGACCTGGAACCAATGACCATCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTC	1580								
QY	1737	ATGCTGAAGAGCAGCAGAGAGAGAAATTCATGAGAAAGGAGCAACTGGCATTCGACG	1796								
DB	1581	ATGCTGAAGAGCAGCAGAGAGAGAAATTCATGAGAAAGGAGCAACTGGCATTCGACG	1640								
QY	1797	TGGCAGTTCTGCTGAAAAGAGAAATGATGCTTTTGAAGACGGAGGAGGAGCAGCTCTTGATGG	1856								
DB	1641	TGGCAGTTCTGCTGAAAAGAGAAATGATGCTTTTGAAGACGGAGGAGGAGCAGCTCTTGATGG	1700								
QY	1857	AGATCAGAGTCTGTCATCGGGCGAGAAACAAGTGACTCTGACAGCAGGCTTACCTTGTTC	1916								
DB	1701	AGATCAGAGTCTGTCATCGGGCGAGAAACAAGTGACTCTGACCAAGGCTTACCTTGTTC	1760								
QY	1917	AAAGAGGAGCTGAGACACAGGAGCTGGCGGCAACAGCGGAATATCCGATTCTCTGCC	1976								
DB	1761	AAAGAGGAGCTGAGACACAGGAGCTGGCGGCAACAGCGGAATATCCGATTCTCTGCC	1820								
QY	1977	CCAAAGTGGAGAGGTTCTGCTGACATAGACACAGCTTACAGATTCACTGATGGATTGCA	2036								
DB	1821	CCAAAGTGGAGAGGTTCTGCTGACATAGACACAGCTTACAGATTCACTGATGGATTGCA	1880								
QY	2037	TCATTTAAAGTGTGATGATATCACTCCCAAACTGTTGGT	2077								
DB	1881	TCATTTAAAGTGTGATGATATCACTCCCAAACTGTTGGT	1921								

RESULT 10

US-10-956-157-5027
; Sequence 5027, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:

; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5027

; LENGTH: 2296

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-956-157-5027

Query Match 85.6%; Score 1777.8; DB 10; Length 2296;

Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1779; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	297	CAGCAACTTCTGCAATGCCATCAACTCTCAGCTGCTCCTCAGTGAAGAGGAGGACAGCC	356
Db	132	CAGCACTTCTGCAATGCCATCAACTCTCAGCTGCTCCTCAGTGAAGAGGAGGACAGCC	191
Qy	357	CCAGTGAAGCACAGGAATAAGACCCGCCCACTGGGCCACCCCAAACTGGACACGTTTA	416
Db	192	CCAGTGAAGCACAGGAATAAGACCCGCCCACTGGGCCACCCCAAACTGGACACGTTTA	251
Qy	417	CCCCGGAGGAGCTGCTCAGCAGATGAAGAGCTCTGACCAAGAACCCACGCTGAAG	476
Db	252	CCCCGGAGGAGCTGCTCAGCAGATGAAGAGCTCTGACCGAGAACCCACGCTGAAG	311
Qy	477	AAGCCATGAACCTAATAATCAAGCCATGAAGAGGAGATTTGAGGAGCTTTGGCCCTGGA	536
Db	312	AAGCCATGAGCTAATAATCAAGCCATGAAGAGGAGATTTGAGGAGCTTTGGCCCTGGA	371
Qy	537	CAGAGAAACAGAGGAAGAACCGCCAGTTTGTGATACAGAGCAAGAGCAAAAGAGC	596
Db	372	CAGAGAAACAGAGGAAGAACCGCCAGTTTGTGATACAGAGCAAGAGCAAAAGAGC	431
Qy	597	GTCTAATGGCTTCAGTCATGAGATGAGAAATGAAGGAGAGCTTTGGAATCTAAAG	656
Db	432	GTCTAATGGCTTCAGTCATGAGATGAGAAATGAAGGAGAGCTTTGGAATCTAAAG	491
Qy	657	GGAATCAGAAAGTCTCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAG	716
Db	492	GGAATCAGAAAGTCTCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAAG	551
Qy	717	CGGAGCAGGAAGGAACGACCTCAGGACCCCAAGTGTGAGGCTACAGCAGAGAGGAGC	776
Db	552	CGGAGCAGGAAGGAACGACCTCAGGACCCCAAGTGTGAGGCTACAGCAGAGAGGAGC	611
Qy	777	ACCTGTTGGGATCGTGTCTGAACTGCACTGCACTGAACTCCAGGGCTCTCAGAG	836
Db	612	ACCTGTTGGGATCGTGTCTGAACTGCACTGCACTGAACTCCAGGGCTCTCAGAG	671
Qy	837	ATTCCTTTGTTGAAATTAGATGCTGAAGAGAGCAGAGGCTCAGTAAAGAAATCA	896
Db	672	ATTCCTTTGTTGAAATTAGATGCTGAAGAGAGCAGAGGCTCAGTAAAGAAATCA	731
Qy	897	AGCATGCTCTGGGCCCAACAGAAACAGTCTCCTGCACTGGCAGCGCAATGCTCTAAATATAGGA	956
Db	732	AGCATGCTCTGGGCCCAACAGAAACAGTCTCCTGCACTGGCAGCGCAATGCTCTAAATATAGGA	791
Qy	957	GCAGATCTGAGATGGGGCCCAAGAAATTAATTCGAACATGAGGAGTTAACTGTGAGCCAGC	1016
Db	792	GCAGATCTGAGATGGGGCCCAAGAAATTAATTCGAACATGAGGAGTTAACTGTGAGCCAGC	851
Qy	1017	TCCTGCTGCTGAAGGAGGGAATCAGAGGTTGAGAGACTTGAAGTTGCACTCAAGG	1076
Db	852	TCCTGCTGCTGAAGGAGGGAATCAGAGGTTGAGAGACTTGAAGTTGCACTCAAGG	911
Qy	1077	AGGCCAAGAAAGAGTTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAA	1136
Db	912	AGGCCAAGAAAGAGTTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAA	971
Qy	1137	CCCAGACAGAGGGGAGCACAGAGAAAGAGATGATGAAGAGAAAGGCCCGGAGACTGTTG	1196
Db	972	CCCAGACAGAGGGGAGCACAGAGAAAGAGATGATGAAGAGAAAGGCCCGGAGACTGTTG	1031
Qy	1197	GAAGGAGTGAAGGAGCTGACCTCCAGGAGTACATCTGTTTAAAGAGCTTCAAGAGG	1256
Db	1032	GAAGGAGTGAAGGAGCTGACCTCCAGGAGTACATCTGTTTAAAGAGCTTCAAGAGG	1091
Qy	1257	CTCATACAAAACCTCAGGCAAGCTCAGCTAATGAAGAGAGACTTCAAGAAAAGTGTGAGG	1316
Db	1092	CTCATACAAAACCTCAGGCAAGCTCAGCTAATGAAGAGAGACTTCAAGAAAAGTGTGAGG	1151
Qy	1317	CCCTTGAAAGGAAAAATCTGCAATTCATCAGAGTTGAATGAAGAAAGAGCTGTTT	1376

Db	1152	CCCTTGAAAGGAAAAATTTCTGCAATTTCCATCAGAGTTGAATGAAGAAAGAGCTTGT	1211
Qy	1377	ATACTAACAAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC	1436
Db	1212	ATACTAACAAAAAGTTAGAGCTACAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAC	1271
Qy	1437	AGGCTAAAAACAGAGGATGAAAGTCCAAATTAATCTGTGCTACAGATGACACAAAGC	1496
Db	1272	AGGCTAAAAACAGAGGATGAAAGTCCAAATTAATCTGTGCTACAGATGACACAAAGC	1331
Qy	1497	TTCTTCAAGAACATATAATGCTTGAAGCAATTTAGGAACTTAAAGAAAGAGCTCAG	1556
Db	1332	TTCTTCAAGAACATATAATGCTTGAAGCAATTTAGGAACTTAAAGAAAGAGCTCAG	1391
Qy	1557	AAAAAGTGGACAGGGCAGTGTCTGAAGCACTGAGTGAAGAACTGGAACTGGCAGAGAGG	1616
Db	1392	AAAAAGTGGACAGGGCAGTGTCTGAAGCACTGAGTGAAGAACTGGAACTGGCAGAGAGG	1451
Qy	1617	CTCTGGCTTCCAAACAGCTGCAAAATGGATGAATGAAGCAAACTTTCGCAAGAGAG	1676
Db	1452	CTCTGGCTTCCAAACAGCTGCAAAATGGATGAATGAAGCAAACTTTCGCAAGAGAG	1511
Qy	1677	AGGACCTGGAACCACTGATCCCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTC	1736
Db	1512	AGGACCTGGAACCACTGATCCCTCAGGGCTCAGATGGAAGTTTACTGTTCTGATTTTC	1571
Qy	1737	ATGCTGAAAGAGCAGCGAGAGAGAAATTCATGAGAAAGAGAGCACTGGCATTGCAGC	1796
Db	1572	ATGCTGAAAGAGCAGCGAGAGAGAAATTCATGAGAAAGAGAGCACTGGCATTGCAGC	1631
Qy	1797	TGSCAGTTCTGTGAAGAGAAATGATGCTTTTGAAGACCGGAGGAGGAGCTCTTGTATGG	1856
Db	1632	TGSCAGTTCTGTGTGAAGAGAAATGATGCTTTTGAAGACCGGAGGAGGAGCTCTTGTATGG	1691
Qy	1857	AGATGAGAGTCTGATGGGCGAGAAACAGTGAATCTGACCGAGGCTTACCTTGTTC	1916
Db	1692	AGATGAGAGTCTGATGGGCGAGAAACAGTGAATCTGACCGAGGCTTACCTTGTTC	1751
Qy	1917	AAAGAGGAGCTGAGGACAGGGACTGGCGGCAACAGCGGAATATTCGGAATTCATTCTGTC	1976
Db	1752	AAAGAGGAGCTGAGGACAGGGACTGGCGGCAACAGCGGAATATTCGGAATTCATTCTGTC	1811
Qy	1977	CCAAGTGTGAGAGGTTCTGCTGACATAGACAGTTTACAGATTCACGTGATGATGCA	2036
Db	1812	CCAAGTGTGAGAGGTTCTGCTGACATAGACAGTTTACAGATTCACGTGATGATGCA	1871
Qy	2037	TCATTTAAGTTGATGATACCTCCCAAACTGTTGCT	2077
Db	1872	TCATTTAAGTTGATGATACCTCCCAAACTGTTGCT	1912

RESULT 11
US-10-609-133-13
; Sequence 13, Application US/10609133
; Publication No. US20040191798A1
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut Health Center
; APPLICANT: Sarfarazi, Mansoor
; APPLICANT: Rezaie, Tayebeh
; APPLICANT: Child, Anne H
; TITLE OF INVENTION: Optineurin and Glaucoma
; FILE REFERENCE: UCT-0046P2
; CURRENT APPLICATION NUMBER: US/10/609,133
; PRIOR FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: 10,281,457
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 10,090,118
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 10,060,981
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: 60/344,754
; PRIOR FILING DATE: 2001-12-24
; NUMBER OF SEQ ID NOS: 18

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 2076
; TYPE: DNA
; ORGANISM: Macaca mulatta
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AY228374
; DATABASE ENTRY DATE: 2002-03-30
; RELEVANT RESIDUES: (1)..(2076)
US-10-609-133-13

Query Match      84.2%; Score 1747.8; DB 9; Length 2076;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1912; Conservative 0; Mismatches 77; Indels 88; Gaps 3;

QY 1 ATCCCGGTGGGAGTTCTCTCCAGGCGGCAACGATGCGGAGGAAAACAGTACCTTGAGCGA 60
D 1 ATCCCGGTGGGAGTTCTCTCCAGGCGGCAACGATGCGGAGGAAAACAGTACCTTGAGCGA 60
QY 61 AGCCAAAGCGGCGGAGGTGTGGCTTTGATAGCTGGTGTGCCACTTCTCTGGCTTGA 120
D 61 AGCCAAAGCGGCGGAGGTGTGGCTTTGATAGCTGGTGTGCCACTTCTCTGGCTTGA 120
QY 121 TGAGCGGTAGCGCTCTGTAAACCCAACTTCTCACTTTGAAACAGCTGCGTGGTTCAGC 180
D 121 TGAGCTGTAGCGCTCTGTAAACCCAGCTTCTCACTAT-AAAACAGCTGCGTGGTTCAGC 179
QY 181 ATTAATGAAGATTAGTCAGTGACAGGCGTGGTGTGCTGAGTCCGCGACATAGAAGATCAA 240
D 180 GTTAATGAAGATTAGTCAGTGACAGGCGTGGTGTGCGGAGTCCACACATA----- 229
QY 241 AATGTCCAAAATGTAATGGAGAGAAAGTGGSCAACTTTTGGAGTGACTTTTCCACAGG 300
D 230 -----G 230
QY 301 AACTTCTGCAATGTCCATCAACCTCTCAGCTGCCTCACTGAAAGAGGAGACAGCCCCAG 360
D 231 AACTTCTGCAATGTCCATCAACCTCTCAGCTGCCTCACTGAAAGGGGACAGCCCCAG 290
QY 361 TGAAGACAGGAAATGGACCCCCCACTTGGCCCAACCAACCTTGGACAGCTTTACCCC 420
D 291 TGAAGACAGGAAATGGACCCCCCACTTGGCCCAACCAACCTTGGACACATTCACCCC 350
QY 421 GGAGGAGCTCTGCAGCAGATGAAAGAGCTTCTGACCAAGAACACACAGCTGAAAGAACG 480
D 351 GGAGGAGCTCTGCAGCAGATGAAAGAGCTTCTGACCGAGAACACACAGCTGAAAGAACG 410
QY 481 CATGAAGCTAAATAATCAAGCCATGAAAGGAGATTGAGGAGCTTTCCGCTTGGACAGA 540
D 411 CATGAAGCTAAATAACCAAGCCATGAAAGGCGGTTTGGAGAGCTTTCCGCTTGGACAGA 470
QY 541 GAAACAGAGGAGAACGCCAGTTTGTGATACAGAGCAAGAGCAAGAGCGTCT 600
D 471 GAAACAGAGGAGAACGCCAGTTTGTGATACAGAGCAAGAGCAAGAGCGTCT 530
QY 601 AATGGCTTCAGTCATGAGAAATGAGAAATGGAAGGAGAGCTTGGAAATCAAGGGGAA 660
D 531 CATGGCTTCAGTCATGAGAAATGAGAAATGGAAGGAGAGCTTGGAAATCAAGGGGAA 590
QY 661 ATCAGAAAGGTCTATCTGAGGACCCCACTGATGACTCCAGGCTTTCCAGGCGCGAAGCGGA 720
D 591 ATCAGAAAGGTCTATCTGAGGACCCCACTGATGACTCCAGGCTTTCCAGGCGCGAAGCGGA 650
QY 721 GCAGGAAAGGACAGCTCAGGACCCAGGTGGTGGAGCTACAGCGAGAGAGCGACCT 780
D 651 GCAGGAAAGGACAGCTCAGGACCCAGGTGGTGGAGCTACAGCGAGAGAGCGACCT 710
QY 781 GTTGGGCATCTGTCTGAACTGACAGCTCAAGCTGAACTCCAGCGGCTCTCTCAGAAGATTC 840
D 711 GTTGGGCATCTGTCTGAACTGACAGCTCAAACTGAACTCCAGCGGCTCTCTCAGAAGATTC 770
QY 841 CTTTGTGTAATATAGGATGGCTGAAAGAGAGAGAGGAGGAGGATGTAAGAAATCAAGCA 900
D 900 CTTTGTGTAATATAGGATGGCTGAAAGAGAGAGAGGAGGATGTAAGAAATCAAGCA 830
QY 901 TAGTCTCGGCGCCACAGAACAGTCTCCACTGGCAGCGCATTTGTCTAAATATAGGAGCAG 960
D 831 TAGTCTCGGCGCCACAGAACAGTCTCCACTGGGAC-----GAGCAG 872
QY 961 ATCTGCAGATGGGCGCCAAAGAAATTACTTCGAAACATGAGGAGTTAACTGTGAGCAGCTCT 1020
D 873 ATCTGCAGAGGGGCGCCAAAGAAATTACTTCGAAACATGAGGAGTTAACTGTGAGCAGCTCT 932
QY 1021 GCTGTGCTTAAGGGAGGGAATCAGAGGTGGAGAGACTTGAAAGTTGCACTCAAGAGGCG 1080
D 933 GCTGTGCTTAAGAGAGGGAATCAGAGGTGGAGAGACTTGAAATTTGCACTCAAGAGGCG 992
QY 1081 CAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAGTAATCGTCTGAGATTGAAACCCA 1140
D 993 CAAAGAAAGAGTTTCAGATTTTGAAGAAAGAAACAGTAATCGTCTGAGATTGAAACCCA 1052
QY 1141 GACAGAGGGGAGCACAGAGAAAGAGATTCATGAAGAGAAAGCCCGGAGACTGTGGGAAG 1200
D 1053 GACAGAGGGGAGCACAGAGAAAGAGAAACGAGAGAGAGAAAGGCCCAGAGACTGTGGGAAG 1112
QY 1201 CGAAGTGAAGACACTGAAACCTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCA 1260
D 1113 CGAAGTGAAGACACTTGAACCTTCCAGGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCA 1172
QY 1261 TACAAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTCCAGGCCCT 1320
D 1173 TACAAAACTCAGTGAAGCCGAGCTAATGAAGAGAGACTTCAAGAAAAAGTGTCCAGGCCCT 1232
QY 1321 TGAAGGAAAAAATTCGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTGTTATAC 1380
D 1233 TGAAGGAAAAAATTCGCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTGTTATAC 1292
QY 1381 TAAACAAAAGTTAGAGCTTACAGTGGAAAGCATGCTATCAGAAATCAAAATGGAAACAGCG 1440
D 1293 TAAACAAAAGTTAGAGCTTACAGTGGAAAGCATGCTCTCGGAAATCAAAATGGAAACAGCG 1352
QY 1441 TAAACACAGAGATGAAAGTCCAAATTAACCTGCTGCTACAGATGACACACAAACAGCTTCT 1500
D 1353 TAAACACAGAGATGAAAGTCCAAATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1412
QY 1501 TCAGAAACATATAATGCAATTTGAAACCAATTTGAGGAACATAACAGAAAGAGTTCAGAAA 1560
D 1413 TCAGAAACATATAATCATGCAATTTGAAACCAATTTGAGGAACATAACAGAAAGAGTTCAGAAA 1472
QY 1561 AGTGAACAGGCGAGTGTGTAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCT 1620
D 1473 AGTGAACAGGCGAGTGTGTAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCT 1532
QY 1621 GGCTTCCAAAACAGCTGCAAAATGGATGAAATGAAGCAAAACATTTGCCAAGCAGGAAGAGA 1680
D 1533 GGCTTCCAAAACAGCTGCAAAATGGATGAAATGAAGCAAAACATTTGCCAAGCAGGAAGAGA 1592
QY 1681 CCTGAAACCATGACCATCTCTCAGGGCTCAGATGGAAGTTTACTGTCTGTGATTTTCATGC 1740
D 1593 CCTGAAACCATGACCATCTCTCAGGGCTCAGATGGAAGTTTACTGTCTGTGATTTTCATGC 1652
QY 1741 TGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAAAGGAGCAACTGGCAATTCGAGCTGGC 1800
D 1653 TGAAGAGCAGCGAGAGAGAAATTCATGAGGAAAAAGGAGCAACTGGCAATTCGAGCTGGC 1712
QY 1801 AGTTCTGCTGAAAGAGAAATGATGCTTTTGAAGCAAGAGGAGCGGAGTCTCTGATGGAGAT 1860
D 1713 AGTTTCTGCTGAAAGAGAAATGATGCTTTTGAAGCAAGAGGAGCGGAGTCTCTGATGGAGAT 1772
QY 1861 GCAGAGTCTGATGGGGGAGAGAAACAGTGAATCTCAAGCAGGCTTACCTTTGTTCAAG 1920
D 1773 GCAGAGTCTGATGGGGGAGAGAAACAGTGAATCTCAAGCAGGCTTACCTTTGTTCAAG 1832
QY 1921 AGGAGCTGAGGACAGGAGCTGGCGCAACAGCGGAAATATTCGATTTTCTTCGCCCCAA 1980
D 1833 AGGAACTGAGGACAGGAGCTGGCGCAACAGCGGAAATATTCGATTTTCTTCGCCCCAA 1892
```


Db	1621	GACAGGAGCTGGCGGCAACAGCGGAATATTCCGATTCAITTCCTGCCCCCAAGTGTGGAGAG	1680
Qy	1991	GTTCCTGCTGACATGACACAGCTTACAGAGTTTCAGGTGATCGATTGCAATCAATTTAA	2044
Db	1681	GTTCCTGCTGACATGACACAGCTTACAGAGTTTCAGGTGATCGATTGCAATCAATTTAA	1734
RESULT 13			
US-10-136-728-71			
; Sequence 71, Application US/10136728			
; Publication No. US20030236188A1			
; GENERAL INFORMATION:			
; APPLICANT: Spytek, Kimberly A.			
; APPLICANT: Li, Li			
; APPLICANT: Edinger, Shlomit R.			
; APPLICANT: Stone, David J.			
; APPLICANT: Guo, Xiaojia			
; APPLICANT: Anderson, David W.			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Gerlach, Valerie L.			
; APPLICANT: Taupier, Raymond J.			
; APPLICANT: Pena, Carol E.A.			
; APPLICANT: Padigaru, Muralidhara			
; APPLICANT: Kekuda, Ramesh			
; APPLICANT: Gorman, Linda			
; APPLICANT: Zerhusen, Bryan D.			
; APPLICANT: Smithson, Glenda			
; APPLICANT: Macdougall, John R.			
; APPLICANT: Mezes, Peter S.			
; APPLICANT: Perman, John A.			
; APPLICANT: Zhong, Mei			
; TITLE OF INVENTION: No. US20030236188A1el Human Proteins, Polynucleotides Encoding Th			
; FILE OF INVENTION: The Same			
; FILE REFERENCE: 21402-347 D (Cura 647 Other)			
; CURRENT APPLICATION NUMBER: US/10/136,728			
; CURRENT FILING DATE: 2002-05-01			
; PRIOR APPLICATION NUMBER: 60/288,395			
; PRIOR FILING DATE: 2001-05-03			
; PRIOR APPLICATION NUMBER: 60/289,087			
; PRIOR FILING DATE: 2001-05-07			
; PRIOR APPLICATION NUMBER: 60/289,619			
; PRIOR FILING DATE: 2001-05-08			
; PRIOR APPLICATION NUMBER: 60/289,818			
; PRIOR FILING DATE: 2001-05-09			
; PRIOR APPLICATION NUMBER: 60/289,817			
; PRIOR FILING DATE: 2001-05-09			
; PRIOR APPLICATION NUMBER: 60/290,194			
; PRIOR FILING DATE: 2001-05-11			
; PRIOR APPLICATION NUMBER: 60/290,753			
; PRIOR FILING DATE: 2001-05-14			
; PRIOR APPLICATION NUMBER: 60/291,189			
; PRIOR FILING DATE: 2001-05-15			
; PRIOR APPLICATION NUMBER: 60/292,374			
; PRIOR FILING DATE: 2001-05-21			
; PRIOR APPLICATION NUMBER: 60/293,107			
; PRIOR FILING DATE: 2001-05-23			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 132			
; SEQ ID NO 71			
; LENGTH: 1908			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (31)..(1714)			
US-10-136-728-71			
Query Match 81.3%; Score 1688.8; DB 7; Length 1908;			
Best Local Similarity 99.6%; Pred. No. 0;			
Matches 1693; Conservative			
; Indels 0; Gaps 0;			
Qy	378	GACCCCCACCTGGCGCCACCCAAACCTGGACACGTTTACCCCGGAGGAGCTGTCGACG	437


```
Qy 1518 CATTGAAACAATTGAGGAATAACAAGAAAGAGTGCAGAAAAAGTGGACAGGCGAGTGC 1577
Db 1190 CATTGAAAAAATAATTGAGGAATAACAAGAAAGAGTGCAGAAAAAGTGGACAGGCGAGTGC 1249
Qy 1578 TGAAGGAACAGTGAAGAACTGGAACTGGCAGAGAGAGGCTCTGGCTTCCAAAAGTGC 1637
Db 1250 TGAAGGAACAGTGAAGAACTGGAACTGGCAGAGAGAGGCTCTGGCTTCCAAAAGTGC 1309
Qy 1638 AAATGGATGAATGAAGCAAAACCAATTCGCAAGCAGGAAGAGGACCTGGAAACCATGACCA 1697
Db 1310 AAATGGATGAATGAAGCAAAACCAATTCGCAAGCAGGAAGAGGACCTGGAAACCATGACCA 1369
Qy 1698 TCCTCAGGCTCAGATGAAGTTTACTGTTCTGTATTTTCTGATTTTCTGATGAAAGAGCAGAG 1757
Db 1370 TCCTCAGGCTCAGATGAAGTTTACTGTTCTGTATTTTCTGATTTTCTGATGAAAGAGCAGAG 1429
Qy 1758 AGAAATTCATGAGGAAAGAGCAACTGGCAATTCGAGCTGGCAGTCTCTGCTGAAAGAGA 1817
Db 1430 AGAAATTCATGAGGAAAGAGCAACTGGCAATTCGAGCTGGCAGTCTCTGCTGAAAGAGA 1489
Qy 1818 ATGATGCTTTGAAAGCGAGGAGGAGTCTTTGATGGAGATGCAGAGTCTCATGGGG 1877
Db 1490 ATGATGCTTTGAAAGCGAGGAGGAGTCTTTGATGGAGATGCAGAGTCTCATGGGG 1549
Qy 1878 CGAACAAGTGAATCTGACAGCAGAGGCTTACCTTGTTCGAAAGAGGAGTGAAGCAGGG 1937
Db 1550 CGAACAAGTGAATCTGACAGCAGAGGCTTACCTTGTTCGAAAGAGGAGTGAAGCAGGG 1609
Qy 1938 ACTGGCGCACAGCGGAATATTCGATTCATTCCTGCCCCCAAGTGGAGAGTTCCTGC 1997
Db 1610 ACTGGCGCACAGCGGAATATTCGATTCATTCCTGCCCCCAAGTGGAGAGTTCCTGC 1669
Qy 1998 CTGACATAGACACCTTACAGATTCACTGATGATTCATTTAAGTGTGATGATC 2057
Db 1670 CTGACATAGACACCTTACAGATTCACTGATGATTCATTTAAGTGTGATGATC 1729
Qy 2058 ACTTCCCCAAACTGTGGT 2077
Db 1730 ACTTCCCCAAACTGTGGT 1749
```

RESULT 14

```
US-10-136-728-69
; Sequence 69, Application US/10136728
; Publication No. US20030236188A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, Kimberly A.
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Anderson, David W.
; APPLICANT: Patturajan, Meera
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Smithson, Glenda
; APPLICANT: MacDougall, John R.
; APPLICANT: Mezes, Peter S.
; APPLICANT: Perman, John A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: No. US20030236188A1 Human Proteins, Polynucleotides Encoding TH
; FILE OF INVENTION: The Same
; FILE REFERENCE: 21402-347 D (Cura 647 Other)
; CURRENT APPLICATION NUMBER: US/10/136,728
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/288,395
; PRIOR FILING DATE: 2001-05-03
```

```
; PRIOR APPLICATION NUMBER: 60/289,087
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 60/289,619
; PRIOR FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: 60/289,818
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: 60/289,817
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: 60/290,194
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/290,753
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/291,189
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/292,374
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/293,107
; PRIOR FILING DATE: 2001-05-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 69
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)..(1714)
US-10-136-728-69
```

```
Query Match 79.3%; Score 1646.6; DB 7; Length 1908;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 7; Indels 18; Gaps 1;
```

```
Qy 379 ACCCCCCCCTGGCCACCCCAACCTGGACAGCTTTACCCCGGAGGAGCTGTCGACGA 438
Db 69 ATCTCTCCCACTGGATCTCCCAAACTGGACACATTTACCCCGGAGGAGCTGTCGACGA 128
Qy 439 GATGAAAGAGCTCTCTGACCAAGAAACCAACAGCTGAAAGAGCCATGAAGCTAAATATCA 498
Db 129 GATGAAAGAGCTCTCTGACCGAGAACCAACAGCTGAAAGAGCCATGAAGCTAAATATCA 188
Qy 499 AGCCATGAAAGGAGATTTGAGGAGCTTTCGGCTCGACAGAGAAAACAGAAAGAAACG 558
Db 189 AGCCATGAAAGGAGATTTGAGGAGCTTTCGGCTCGACAGAGAAAACAGAAAGAAACG 248
Qy 559 CCAGTTTTTTGATACAGAGCAAGCAAGCAAGAGCGCTTAATGCGCTTGAGTCATGA 618
Db 249 CCAGTTTTTTGATACAGAGCAAGCAAGCAAGAGCGCTTAATGCGCTTGAGTCATGA 308
Qy 619 GAATGAGAAATTGAAGGAAAGAGCTTGGAAAACTTAAAGGGAAATCAGAAAGGTCACTGA 678
Db 309 GAATGAGAAATTGAAGGAAAGAGCTTGGAAAACTTAAAGGGAAATCAGAAAGGTCACTGA 368
Qy 679 GGACCCCACTGATGATCTCCAGGCTTCCAGGGCCGAGCGGAGCAGAAAAGACAGCT 738
Db 369 GGACCCCACTGATGATCTCCAGGCTTCCAGGGCCGAGCGGAGCAGAAAAGACAGCT 428
Qy 739 CAGGACCCAGTGGTGAGGCTACAGCAGAGAGGAGCAGCTGTTGGGCATCGTGTCTGA 798
Db 429 CAGGACCCAGTGGTGAGGCTACAGCAGAGAGGAGCAGCTGTTGGGCATCGTGTCTGA 488
Qy 799 ACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAAAGATTCTCTTTGTTGAAATTAGGAT 858
Db 489 ACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAAAGATTCTCTTTGTTGAAATTAGGAT 548
Qy 859 GCGTGAAGGAGAGCAAGAGGCTCAGTAAAGAAATCAAGCATAGTCTCTGGGCCACGAG 918
Db 549 GCGTGAAGGAGAGCAAGAGGCTCAGTAAAGAAATCAAGCATAGTCTCTGGGCCACGAG 608
Qy 919 AACAGTCTCACTGGCAGCGGATTTGCTAAATATAGGAGCAGATCTGAGATGGGGCCAA 978
Db 609 AACAGTCTCACTGGCAC-----GAGCAGATCTGCAGATGGGGCCAA 650
```


QY 383 CCCACCTGGCCCAACCCAAACCTGGACACAGTTTACCCCGGAGGAGCTGCTGCACAGATG 442
Db 223 TCCAATATGGTTTACCCCGAGCTGGACACATTCACCCCTGAGGAGCTGCTGCACAAATG 282
QY 443 AAGAGCTCTCTGACCAAGAACCCAGCTGAAGACCATGAAGCTTAATTAATCAAGCC 502
Db 283 AAGGACTCTCTGGTTGAGAACCAACAGCTGAAGACCATGAAGCTTAATTAATCAAGCT 342
QY 503 ATGAAGGGAGATTGAGGAGCTTTCCGCTGGACAGAGAAACAGAAAGGAGAAAGCCAG 562
Db 343 ATGAAGGGCGATTGAGGAGCTTCCGCTGGACAGAGAAAGGAGAGAGCGCTG 402
QY 563 TTTTTCAGATACAGAGCAAGAAAGCAAGAGCGTCTAATGCGCTTGAAGTATGAGAT 622
Db 403 TTGTTTGAAGATGCAAGCAAGAGAGTTAAGGAGCGCTTGAAGCGCTTGAAGTATG 462
QY 623 GAGAAATTTGAAGGAGAGCTTTGAGAACTTAAAGGGAAATCAGAAAGGTCATCTGAGGAC 682
Db 463 GAGAGGCTGAAGGAGAGCTTTGAGAAATTTCAAGAGAAATCAGAAAGGCTTGAAGAC 522
QY 683 CCCACTGATGACTCCAGGCTTCCAG-----GGCC 712
Db 523 CTCACGCTGGCTACAGTATCCAGAGCTTGGAGGAGGAGTGGAGAGCTGAAGACC 582
QY 713 GAAGCGAGAGCAAGAAAGGACAGCTCAGGACCCAGGTGGTGAGGCTACAGAGAGAG 772
Db 583 CAGGTGAGAGGAGGAGTGGAGCATCTGAAGATCCAGGTGATGCGCTTCCGCGCTGAAAG 642
QY 773 GCAGACCTGTTGGGCATCTGCTGAACTGCACTCAGCTCAGCTCAGCTCAGCGCTCTCA 832
Db 643 GCAGACCTGTTGGGCATCTGCTGAACTGCACTCAGCTCAGCTCAGCTCAGCGCTCTCG 702
QY 833 GAAGATTCCTTTGTTGAAATTTAGGATGGCTGAAGGAGAGCAGAAAGGTCAGTAAAGAA 892
Db 703 GAAGACTCTCTGTTGAGATCAGGATGACCGAGGAGAGACTGAAGGGGCAATGAGAG 762
QY 893 ATCAGCATATGCTCGGGCCACAGAGACAGTCTCCACTGCGCAGGCAATGTCTAAATAT 952
Db 763 ATGAAGAACTGCTTACACCCACAGAAACAGACCCCATCAGC-----804
QY 953 AGGAGCAGATCTGAGATGGGCCCAAGAAATTACTTCGAAATGAGGAGTTAACTGTGAGC 1012
Db 805 TTGAGCAACTGTACAGAGGATGCCAGGAGTTGTCGGAGTTTGAAGAACTGACTGTGAGC 864
QY 1013 CAGCTCTGCTGCTCCTAAGGAGGAGGAATCAGAAAGTGGAGACTTTGAAGTTGCACTC 1072
Db 865 CAGCTCTGCTTTCCTTAAGGAGGAGAAACCAAAAGTGGAGACTTTGAAGTGGCCCTC 924
QY 1073 AAGAGGCCAAAGAGAGTTTCAGATTTTGAAGAAAGAAACAGATATCGTTCTGAGATT 1132
Db 925 AGAGAGCCAAAGAGAAATTTCAATTTTGAAGAAAGAAACCAATGGCCATTTCTTACT 984
QY 1133 GAAACCCAGACAGAGGGGAGCAGAGAAAGAGATGATGAGAAAGGCCCGGAGACT 1192
Db 985 GAGAGCAGACAGCGGAGGAGCAGACAGAGAGA---AGGAGGACAAAGGCCCAAGAGAT 1041
QY 1193 GTTGAAGCGAAGTGAAGCACTGAACTCCAGTGAATCTCTGTTTAAAGGAGCTTCAA 1252
Db 1042 GTTGAAGCGAAGTGAAGAACTGAGCAATTCAGTGAATCTCTGTTTAAAGGAGCTTCAA 1101
QY 1253 GAGCTCATACAAACTCAGCGAAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGTGT 1312
Db 1102 GAGGACACACAAACTCAGTGAAGCTGAGCTGATGAAGAGAGACTTCAAGAAAGTGT 1161
QY 1313 CAGGCTTGAAGGAGAAATTCGCAATTCATCAGAGTTGAATGAAGAGAGAGCTT 1372
Db 1162 CAGGCTGAGAGGAGAACTCTGCAACACCATCAGAGCTGATGAAGAGAGAGCTC 1221
QY 1373 GTTTATTAACAAAAAGTTAGAGCTACAAAGTGAAGAGATGCTTATCAGAAATCAAAATG 1432
Db 1222 GTTTACGTAACAAAGATTTAGAGCTCAGGTGAGAGCATGCGCTCCGAAATCAAGATG 1281
QY 1433 GAACAGGCTAAAAACAGAGGATGAAAGTCCAAATTTAACTGTGCTACAGATGACACAAAC 1492

Db 1282 GAGCAGGCCAAGACAGAGGAGGAGAGTCTCAGGTTAGCCACTCTGCAGGCAACTCACAAC 1341
QY 1493 AAGCTTCTTCAAGAACATAATTAATGCAATTAAGAACATTTGAGGAATTAAGAACAAAGAG 1552
Db 1342 AAGCTTCTTCAAGAACATAATTAAGGCACTGAAGAACATTTGAAGAACATTAAGAACAAAGAG 1401
QY 1553 TCAGAAAGAGTGGACAGGCGAGTCTGCTGAAGGAACTGAGTGAAGAACTGGAACTGGCAGAG 1612
Db 1402 GCAGAAAGGTGGACAAAGATGTTGCTGCAGGAGCTCAGCAGAGAGCTGGAGCTGGCAGAG 1461
QY 1613 AAGCTCTGCTTCCAAACAGCTGCAGAAATGGAATGAAGCAAAACCAATTCAGCAAGAG 1672
Db 1462 CAGGCTCTGCACTCCAAACAGCTCCAGATGATGAGATGAAGCAGAGCGCTCGCTAAGCAG 1521
QY 1673 GAAGAGACCTGGAAACCATGACCATCTCAGGCTCAGATGGAAGTTTACTGTTCTGAT 1732
Db 1522 GAGGAAGCTGGAGACCATGGCCCTCTCAGGCTCAGATGGAAGTTTACTGCTCAGAT 1581
QY 1733 TTTTCATGCTGAAGAGCAGCAGAGAGAAATTCATGAGGAAAGGAGCAACTGGCATTTG 1792
Db 1582 TTTTCAGCTGAGAGAGCAGCAAGAGAGATTCATGAAGAAAGGAGCAGCTGGCCTTG 1641
QY 1793 CAGCTGCACTTCTGCTGAAGAGATGATGCTTTTGAAGACGAGGC---AGCAGTTC 1849
Db 1642 CAGCTCGCATTTTGTCTGAAGAGAAACAATGACATTTGAAGAGGAGGAGCAGTACAGTTC 1701
QY 1850 TTGATGAGATGAGAGTCTGCTGAGGCGCAGAACAGTCTGCTGACACGAGCGCTTAC 1909
Db 1702 CTGATGAAATGAGTCCGACACGGGGCAAGAACAGTCTGCTGACACGAGCAGCTTAC 1761
QY 1910 CTGTTTCAAGAGGAGCTGAGGAGCAG-----GGACTGGCGGCAACAGCGGAATATT 1960
Db 1762 CTGTTTCAAGAGGAGCGGAGGAGCAGGAGCTGGCAGACCGGSCAGCAGCTCGCAGTATT 1821
QY 1961 CCGATTCATTCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATGACATGACAGTTCAGATT 2020
Db 1822 CCGATTCATTCTGCCCCCAAGTGTGGAGAGGTTCTGCTGACATGACATGACAGTTCAGATT 1881
QY 2021 CAGCTGATGATGTCATCATTTAAGTGTGATGATGATCATCCCTCCCAAACTGTTGTT 2077
Db 1882 CATGTGATGACTGTCATCATTTGAGTGT---TCTCTCAGTCCCAAAAGCTCTTGGT 1935

RESULT 18
US-10-609-133-15
; Sequence 15, Application US/10609133
; Publication No. US20040191798A1
; GENERAL INFORMATION:
; APPLICANT: University of Connecticut Health Center
; APPLICANT: Sarfarazi, Mansoor
; APPLICANT: Rezaie, Tayebbeh
; APPLICANT: Child, Anne H
; TITLE OF INVENTION: Optineurin and Glaucoma
; FILE REFERENCE: UCT-0046P2
; CURRENT APPLICATION NUMBER: US/10/609,133
; CURRENT FILING DATE: 2003-06-26
; PRIOR FILING DATE: 2002-10-25
; PRIOR FILING DATE: 2002-10-25
; PRIOR FILING DATE: 2002-10-25
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2002-01-30
; PRIOR FILING DATE: 2002-01-30
; PRIOR FILING DATE: 2001-12-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 1787
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_145081

; DATABASE ENTRY DATE: 2003-04-06
; RELEVANT RESIDUES: (1)..(1787)
US-10-609-133-15

Query Match 50.9%; Score 1057; DB 9; Length 1787;
Best Local Similarity 77.7%; Pred. No. 9.7e-282;
Matches 1396; Conservative 0; Mismatches 330; Indels 71; Gaps 7;

QY 297 CAGGAATCTTCGAATGTCCTCACTCAACCTCTCAGCTGCTCACTCACTGAAAGGAGGACAGCC 356
DB 13 CCGGACCTGTTACATGTTCCCATCAACCTCTGAGCTGCTGCTGACTGAGAAAGGGGACAGCT 72
QY 357 CCAGTGAAGACGACGAAATGGAACCCGCCACCTGGCCACCCCAACCACTGGACACGTTTA 416
DB 73 CCGTGTGAGACCCAGGAAATGACCTCCAAATATGTTTCAACCCCAACCTGGACACATTCA 132
QY 417 CCGGAGAGAGCTGTGACGACGATGAAGAGCTCTGACCAAGAACCAACGAGCTGAAG 476
DB 133 CTCCTGAGGAGCTGTGACGACCAATGAAGAACTCTGCTGTCGAGAACCAACGAGCTGAAG 192
QY 477 AAGCCATGAAGCTTAATATCAAGCCATCAAGGAGATTTGAGGAGCTTTCCGCTGGA 536
DB 193 AAGCCATGAAGCTTAATATCAAGCTATGAAGGAGCTTTGAGGAGCTTTCCGCTGGA 252
QY 537 CAGAGAAA CAGAGGAAGAACCGCAGTCTTTTGAATACAGAGCAAGAAAGCAAAAGAGC 596
DB 253 CAGAGAGGCAAGAGGAAGAACCGCAGTCTTTTGAATACAGAGCAAGAAAGGCTTAAGGAGC 312
QY 597 GTCTAATGCGCTTGAGTCATGAGATGAGAAATGGAAGAAAGCTTGGAACCTTAAG 656
DB 313 GCCTCAAGGCCCTGAGTCATGAAATGAGAGGCTGAAGGAGAGCTTTGGAACCTTAAG 372
QY 657 GGAATTCAGAAAGCTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAGGGCG --- 713
DB 373 AGAATTCAGAAAGGCCATTTGAAGACATCAGAGTTAGTGCAGGTTTCCAGAACCGACT 432
QY 714 -----AAGCGGAGCAGGAAAGGACCGAGCTCAGG 742
DB 433 TGGAGCAGGAAGTGGGAGCAACTGAAGAGCGAGGTGAGAGCAAGTGGAGCATCTGAAG 492
QY 743 ACCAGGTGGTAGGCTTACAAGC ---AGAGAGGCGAGACCTGTGTGGGCATGTGTCTGAA 799
DB 493 ATCCAGGTGAGGCGCTTCAGGCTTGAGAAAGCGACCTTGCTGGGGCAATTCGTCTCAGA 552
QY 800 CTGACGCTCAAGCTGAACCTCAGCGGCTCTCAGAGAGATTCTTCTGTAATAGGATG 859
DB 553 CTGACGCTCAAGCTCAACTCCGCGGCTCTCAGAAAGACTCTCTCTGTGGAGA -CAGGATG 611
QY 860 GCTGAAGGAGAGCAGAGGGTCAAGTAAAGAAATCAAGCATAGTCTCGGGCCACGAGA 919
DB 612 ACTGAAGGAGAGCGGAGGCGCAATGAAGGAGATGAGGAACAGCGCTGGACCCCAAGS 671
QY 920 ACAGTCTCACTGGACCGGATTTGCTTAATATAGGAGCAGATCTCAGATGGGGCAAG 979
DB 672 ACAGACTCCATCATCTGG -----GCAAAATGTACAGAGGACGCGCAGS 713
QY 980 AATTACTTCGAACATCAGAGTTAACTGTGAGCCAGCTCTCTGTGCTGTAGGGAGGS 1039
DB 714 ACTTGTGTGGAGTTTGAGGAACATGACTGTGAGGCAACTCTCTGTCTTGTGCTCAGGGAAGGA 773
QY 1040 AATCAGAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAAGAAAGAGTTTTCAGAT 1099
DB 774 AACCAGAAAGGTGGAGAGCTCGAGATCGCGCTCAGAGAGGCCAAAGAAAGAAATTTTCAGAT 833
QY 1100 TTTGAAAGAAAACAAGTAACTGTTCTGAGATTGAAACCCAGACAGAGGGGAGCAGAG 1159
DB 834 TTTGAAAGAAAACAAGTAACTGTTCTGCGATTGAAACCCAGACAGAGGGGAGCAGACAA 893
QY 1160 AAGAGAAATGATCAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGAAGCACTGAAC 1219
DB 894 AAG ---AAGAGGAGGACAAAGAACCCAGAGAGTGTGGGAATCGAGTGGAAATCTTGAAC 950
QY 1220 CTCAGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATCAAAAATCAGCGAAGCT 1279

DB 951 GTTCAAGTGGCTCTCTCTGTTTAAAGGGCTTCAAGGGCGCACAAAAGCTCAGTGAGGCC 1010
QY 1280 GAGCTAATGAAGAGAGCTTCAAGAAAGTCTCAAGCCCTTGAAGAGGAAATTTCTGCA 1339
DB 1011 GAGCTGATGAAGAGAGACTTCAAGAAAGTGTCAAGGCTCTGGAAGGAGAACTCTGCA 1070
QY 1340 ATTCCATCAGAGCTGAATGAAGCAAGAGCTTGTATTACTAAACAAAAGTTAGAGCTA 1399
DB 1071 ACCCATCGGAGCTGAATGAAGCAAGAGCTCTGTTACAGTAACTGGAAGTTAGAGCTG 1130
QY 1400 CAAAGTGAAGCATGCTATCAGAATCAAAATGGAACAGGCTAAACAGAGGATGAAAG 1459
DB 1131 CAGGTGGAGAGCATGCGCTCGGAATCAAGATGGAGCGCCAAAGACAGAGGAGGAGAA 1190
QY 1460 TCCAAATTAATCTGTCTACAGATGACACACAAAGCTTCTTCAAGAAACATATAATGCA 1519
DB 1191 TCCAGGTTAGCCACTCTTCAGGCAACACACGACAGCTCTTTCAGGAACCAATAAAGCT 1250
QY 1520 TTGAAAAACAATGAGGAACCTAAAGAAAGAGTCAAGAAAAAGTGAAGGGCAGTGCTG 1579
DB 1251 CTGAGAACATTTGAAGAACTAACCAACACAGGCAAGAAAGTGAAGGTCAGCTG 1310
QY 1580 AAGGAATGAGTGAAGAACTGGAACCTGGCAGAGAGGCTCTGGCTTCCAAACAGCTGCAA 1639
DB 1311 CAGGAGCTCAGCGAGAGAGCTGGAGCTGGCGGAGCAGGCTCTGGCGTCCAAGCAGCTCCAG 1370
QY 1640 ATGGATGAATCAAGCAACCACTTCCCAAGCAGGAGAGGAGCTTGAACCATGACCATC 1699
DB 1371 ATGGATGAGATGAAGCAGACCATCCCAAGCAGGAGGAGGAGCTTGAAGACCATGGCGCTC 1430
QY 1700 CTCAGGCTCAGATGGAAGTCTTACTGTTCTGATTTTCTGCTGATAAGAGCAGCGAGAG 1759
DB 1431 CTCAGGCTCAGATGGAAGTCTTACTGTTCTGATTTTCTGCTGATAAGAGCAGCGAGAG 1490
QY 1760 AAAATTCATGAGAAAGAGCAACTGCAATTCGAGCTGGCAGTCTCTGCTGAAGAGAAAT 1819
DB 1491 AAGATCCATGAGAAAGAGCAGCTGGCTTTCAGCTCGCCATTTTGTGTAAGAGAAC 1550
QY 1820 GATGCTTTGAGAGCGAGGC ---AGGCAGTCTTGTAGAGATGAGATGAGATGCTCATGG 1876
DB 1551 AATGACTTTGAAGATGAGGAGGAGTGGCAGTCTTGTGTAAGATGAGTCCCGGACGGG 1610
QY 1877 GCGAGAACAAAGTCACTCTGACAGCAGGCTTACCTTTGTTCAAAGAGGAGCTGAGGACAGG 1936
DB 1611 GCGAGAACAGGAGCTCTGACAGCAGGCTTACCTGTTTCAAAGAGGAGCCAGGACATG 1670
QY 1937 GACT-----GGCGGCAACAGCGGAATATTCGATTCATTCCTGCCCCCAAGTGTGGA 1987
DB 1671 AGCTGGCAGCATGGGCGAGCGCCCGAGTATTTCCCATTTCACTGCTGCCCCCAAGTGTGG 1730
QY 1988 GAGGTTCTGCTGACATAGACAGCTTACAGATTCAGTGTGATGATTCATTTAA 2044
DB 1731 GAGGCTCTGCGGCAATTCAGACGCTTTCAGATACATGTGTGACTGTATCATCTGA 1787

RESULT 19
US-10-779-543-5625
; Sequence 5625, Application US/10779543
; Publication No. US2005022791A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23

FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 111262
LENGTH: 430
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_32962C.1
US-10-425-115-111262

Query Match 20.4%; Score 423.6; DB 9; Length 430;
Best Local Similarity 99.1%; Pred. No. 2.4e-106;
Matches 426; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1514 AATGCATTGAAACAAATTTAGGAACTAAACAAGAAAGAGTCAGAAAAAGTTGGACAGGGCA 1573
DB 1 AATGCATTGAAACAAATTTAGGAACTAAACAAGAAAGAGTCAGAAAAAGTTGGACAGGGCA 60
QY 1574 GTGCTGAAGAACTGAGTGAAGAACTGGAACCTGSCAGAGAGGCTCTGGCTTCCAAACAG 1633
DB 61 GTGCTGAAGAACTGAGTGAAGAACTGGAACCTGSCAGAGAGGCTCTGGCTTCCAAACAG 120
QY 1634 CTCGAATGATGAATGAAGCAAAACCAATGCCAAGCAGGAAGGACCTGGAAACCATG 1693
DB 121 CTCGAATGATGAATGAAGCAAAACCAATGCCAAGCAGGAAGGACCTGGAAACCATG 180
QY 1694 ACCATCTCAGGCTCAGATGGAAGTTTACTGTCTGATTTTCTGATTTTCTGATGGAAGGAGCG 1753
DB 181 ACCATCTCAGGCTCAGATGGAAGTTTACTGTCTGATTTTCTGATTTTCTGATGGAAGGAGCG 240
QY 1754 AGAGAGAAATTCATGAGGAAAGGAGCAACTGGCAATTCGAGCTGGCAGTTCTGCTGAAA 1813
DB 241 AGAGAGAAATTCATGAGGAAAGGAGCAACTGGCAATTCGAGCTGGCAGTTCTGCTGAAA 300
QY 1814 GAGAATGATGCTTTCGAAGACGAGGAGGAGGAGTCTTGTATGAGATGCAAGTCTGTCAT 1873
DB 301 GAGAATGATGCTTTCGAAGACGAGGAGGAGGAGTCTTGTATGAGATGCAAGTCTGTCAT 360
QY 1874 GGGGCGAGAACAACTGACTCTGACCAAGCAGGCTTACCTCTTCAAGAGGAGCTGAGGAC 1933
DB 361 GGGGCGAGAACAACTGACTCTGACCAAGCAGGCTTACCTCTTCAAGAGGAGCTGAGGAC 420
QY 1934 AGGAGCTGGC 1943
DB 421 AGGAGCTGGC 430

RESULT 23
US-09-960-352-13021
Sequence 13021, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengping
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 13021
LENGTH: 443
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 56-BOVMS1-021-Q1-E1-F8
US-09-960-352-13021

Query Match 17.2%; Score 358.2; DB 3; Length 443;
Best Local Similarity 88.0%; Pred. No. 3.4e-88;

Matches 390; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 1005 CTGTGAGCCAGCTCCTGTGCTAAGGGAAGGAATCAGAAGCTTGAGAGACTTGAAG 1064
DB 1 CTGTGAGCCAGCTCCTGTGCTAAGGGAAGGAATCAGAAGCTTGAGAGACTTGAAG 60
QY 1065 TTGCACTCAAGGAGGCCCAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCGTT 1124
DB 61 TTGCACTCAAGGAGGCCCAAGAAAGAGTTTCAGATTTTGAAGAAACAAAGTAATCGTT 120
QY 1125 CTGAGATTGAAACCCAGACAGAGGGGAGCAGACAGAAAGAAAGTAATGATCAAGAGAAAGGCC 1184
DB 121 CTGAGATTGAAACCCAGACAGAGGGGAGCAGACAGAAAGAAAGTAATGATCAAGAGAAAGGCC 180
QY 1185 CGGAGACTGTTGGAAGCGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAGG 1244
DB 181 CCAAACTATTGGAAGTGAAGTGAACACTGAACCTTCAGGTGACAACTCTGTTTAAGG 240
QY 1245 AGCTTCAAGAGGCTCATACAAAACCTCAGCAAGCTGAGCTTAATGAAGAGAGACTTCAG 1304
DB 241 AGCTTCAAGAGGCTCATACAAAACCTCAGCAAGCTGAGCTTAATGAAGAGAGACTTCAG 300
QY 1305 ABAAGTGTCAAGGCCCTTGAAGGAAATTTCTGCAATTCATCAGAGTTGAATGAAGC 1364
DB 301 ABAAGTGTCAAGGCCCTTGAAGGAAATTTCTGCAATTCATCAGAGTTGAATGAAGC 360
QY 1365 AAGAGCTGTTTATACCAAAAAGTTAGAGCTTACAAGTGAAGAGCATGCTATCAGAAA 1424
DB 361 AAGAGCTGTTTATACCAAAAAGTTAGAGCTTACAAGTGAAGAGCATGCTATCAGAAA 420
QY 1425 TCAAAATGGAACAGGCTTAAACA 1447
DB 421 TCAAAATGGAACAGGCTTAAACA 443

RESULT 24
US-09-960-352-4954
Sequence 4954, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengping
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 4954
LENGTH: 435
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 22-BOVMS1-017-Q1-E1-F5
US-09-960-352-4954

Query Match 16.2%; Score 335.8; DB 3; Length 435;
Best Local Similarity 85.7%; Pred. No. 5.5e-82;
Matches 373; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 956 AGCAGATCTGCAGATGGGCGCAAGAATTTACTTCGAACATGAGGAGTTAACTGTGAGCCAG 1015
DB 1 AGCAATCTGCAGAGGTACCAGGAATTAATGTGGAGTTTCAGGAATTAATCTGTGAGCCAG 60
QY 1016 CTCCTGTGCTTAAGGGAAGGAATCAGAAGTGGAGAGACTTGAAGTTGCACTCAAG 1075
DB 61 CTCCTGTGCTTAAGGGAAGGAATCAGAAGTGGAGAGACTTGAAGTTGCACTCAAG 120
QY 1076 GAGGCCAAAGAGTTTCAGATTTTGAAGAAAGAAACAAAGTAATCTCTGAGATTGAA 1135
DB 121 GAAGCCAAAGAGAAATTTCTGATTTTGAAGAAAGAAAGCAAGGATCATCTGAGACTGAG 180

LOCATION: (322)..(322)
FEATURE:
NAME/KEY: misc feature
LOCATION: (325)..(325)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (327)..(327)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (332)..(332)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (335)..(335)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (340)..(340)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (346)..(346)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (350)..(350)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (357)..(357)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (359)..(361)
OTHER INFORMATION: n is a, c, g, or t
US-11-060-867-55

Query Match 15.9%; Score 330; DB 13; Length 364;
Best Local Similarity 93.0%; Pred. No. 2e-80;
Matches 330; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 979 GAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCTGCTGTGCTTAAGGGAAGG 1038
Db 10 GAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCTGCTGTGCTTAAGGGAAGG 69
QY 1039 GAATCAGAGGTCGAGAGACTTCAAGTTGCACTCAAGGAGCCCAAGAAAGAGTTTCAGA 1098
Db 70 GAATCAGAGGTCGAGAGACTTCAAGTTGCACTCAAGGAGCCCAAGAAAGAGTTTCAGA 129
QY 1099 TTTTGAAGAAAGAAAGTAACTCTGAGATTGAAACCCAGACAGAGGGGAGCACAGA 1158
Db 130 TTTTGAAGAAAGAAAGTAACTCTGAGATTGAAACCCAGACAGAGGGGAGCACAGA 189
QY 1159 GAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTTTGGAAGCGAAGTGGAAAGCACTGAA 1218
Db 190 GAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTTTGGAAGCGAAGTGGAAAGCACTGAA 249
QY 1219 CCTCCAGGTACATCTCTGTTTAAGGAGCTTCAAGAGGCTCATACAAAACCTCAGCGAAGC 1278
Db 250 CCTNCAAGGTACATNTGTTTAAAGGAGCTTNAANAGGCTTNAACAAAACCTNANCAANC 309
QY 1279 TGAGCTAAATGAAGAGAGACTTCAAGAAAGTGTGAGGCCCTTGAAGGAAAT 1333
Db 310 TGAACATGAAGAAAGAACTTNAANAAATGTNANGCCNTTGAAGANNNAAT 364

RESULT 27
US-09-918-995-7614
; Sequence 7614, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 7614
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7614

Query Match 14.3%; Score 296.2; DB 3; Length 432;
Best Local Similarity 99.0%; Pred. No. 5.4e-71;
Matches 298; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 298 AGGAACCTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGGAGGACAGCCC 357
Db 130 AGGAACCTTCTGCAATGTCCCATCAACCTCTCAGCTGCCTCACTGAAAGGAGGACAGCCC 189
QY 358 CAGTGAAGCACAGGAAATGGACCCGCCACCTGGGCCACCCAACTGGACACGTTTAC 417
Db 190 CAGTGAAGCACAGGAAATGGACCCGCCACCTGGGCCACCCAACTGGACACGTTTAC 249
QY 418 CCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTCAGCAGAAAGAGCTCTCAGCAGAAAGAG 477
Db 250 CCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTCAGCAGAAAGAGCTCTCAGCAGAAAGAG 309
QY 478 AGCCATGAAGCTAAATAATCAAGCCATGAAGGAGAGATTTGAGGAGCTTTGGGCTGGAC 537
Db 310 AGCCATGAAGCTAAATAATCAAGCCATGAAGGAGAGATTTGAGGAGCTTTGGGCTGGAC 369
QY 538 AGAAGAACAGAGAGAGAGCCAGTTTGTGATGATACAGAGCAAGCAAGCAAGAGCG 597
Db 370 AGAAGAACAGAGAGAGAGCCAGTTTGTGATGATACAGAGCAAGCAAGCAAGAGCG 429
QY 598 T 598
Db 430 T 430

RESULT 28
US-11-128-061-2706
; Sequence 2706, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2706

```
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-2706

Query Match      14.1%; Score 292.4; DB 16; Length 505;
Best Local Similarity 80.8%; Pred. No. 6.6e-70;
Matches 341; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 289 CTTTTCACAGGAACCTTCGCAATGTCCCATCAACTCTCAGCTGCCTCAGTCAAGAAAGGA 348
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 CTAAGCCGACGGACCTGTATTACATGTCCCATCAACTCGAGCTGCCTGACTGAGAGGG 142
QY 349 GGACAGCCCCAGTGAAGACACAGGAAATGGACCCCCACCTGGCCCCACCCAAACCTGGA 408
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
143 TGACAGCCTCTGTGAGACCCGAGGAAATGAACCCCTAATAAGGTTTCAACCCAGCCTGGA 202
QY 409 CACGTTTACCCCGGAGGAGCTGCTCAGCAGATGAAGAGCTTCCAGCAAGAACCA 468
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
203 CACATTCTCTCTGAGGAGATGCTGCAGCAAAATGAAGGAACCTCTGGTTGAGAACCA 262
QY 469 GCTGAAAGAGCCATGAAGCTAAATCAAGCCATGAAGAGGAGATTTGAGGAGCTTTC 528
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 GCTGAAAGAGCCATGAAGCTAAATCAAGCCATGAAGAGGAGATTTGAGGAGCTTTC 322
QY 529 GGCCTGGACAGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 588
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
323 AGCCTGGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 382
QY 589 AAAAGAGCGTCTAATGGCCCTTGAGTCATGAGAAATGAGAAATGAGAAAGAGCTTGGAAA 648
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
383 TAAGAAGCTCTTGATGACCTTGAGTTATGAAATGAGAACTGAAGGTGAGCTTGGAAA 442
QY 649 ACTAAAGGGAATCGAAGAGTCTCAGAGACCCACTGATGATCAGAGCTCCAGGCTCCAG 708
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 ACTAAGAGAGAGTCTGAGAAAGGCAATTTGAAGGCTCCATATGTGGCTCCAGGCTCCCAT 502
QY 709 GG 710
Db      |||
503 GG 504

RESULT 29
US-11-128-061-6348
; Sequence 6348, Application US/11128061
; Publication No. US2006003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; PRIOR FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6348
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-6348

Query Match      14.1%; Score 292.4; DB 16; Length 505;
Best Local Similarity 80.8%; Pred. No. 6.6e-70;
Matches 341; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
```

```
QY 289 CTTTTCACAGGAACCTTCGCAATGTCCCATCAACTCTCAGCTGCCTCAGTCAAGAAAGGA 348
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 CTAAGCCGACGGACCTGTATTACATGTCCCATCAACTCGAGCTGCCTGACTGAGAGGG 142
QY 349 GGACAGCCCCAGTGAAGACACAGGAAATGGACCCCCACCTGGCCCCACCCAAACCTGGA 408
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
143 TGACAGCCTCTGTGAGACCCGAGGAAATGAACCCCTAATAAGGTTTCAACCCAGCCTGGA 202
QY 409 CACGTTTACCCCGGAGGAGCTGCTCAGCAGATGAAGAGCTTCCAGCAAGAACCA 468
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
203 CACATTCTCTCTGAGGAGATGCTGCAGCAAAATGAAGGAACCTCTGGTTGAGAACCA 262
QY 469 GCTGAAAGAGCCATGAAGCTAAATCAAGCCATGAAGAGGAGATTTGAGGAGCTTTC 528
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
263 GCTGAAAGAGCCATGAAGCTAAATCAAGCCATGAAGAGGAGATTTGAGGAGCTTTC 322
QY 529 GGCCTGGACAGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 588
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
323 AGCCTGGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 382
QY 589 AAAAGAGCGTCTAATGGCCCTTGAGTCATGAGAAATGAGAAATGAGAAAGAGCTTGGAAA 648
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
383 TAAGAAGCTCTTGATGACCTTGAGTTATGAAATGAGAACTGAAGGTGAGCTTGGAAA 442
QY 649 ACTAAAGGGAATCGAAGAGTCTCAGAGACCCACTGATGATCAGAGCTCCAGGCTCCAG 708
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 ACTAAGAGAGAGTCTGAGAAAGGCAATTTGAAGGCTCCATATGTGGCTCCAGGCTCCCAT 502
QY 709 GG 710
Db      |||
503 GG 504

RESULT 30
US-11-128-049-2706
; Sequence 2706, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FO
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2706
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-049-2706

Query Match      14.1%; Score 292.4; DB 16; Length 505;
Best Local Similarity 80.8%; Pred. No. 6.6e-70;
Matches 341; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 289 CTTTTCACAGGAACCTTCGCAATGTCCCATCAACTCTCAGCTGCCTCAGTCAAGAAAGGA 348
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 CTAAGCCGACGGACCTGTATTACATGTCCCATCAACTCGAGCTGCCTGACTGAGAGGG 142
QY 349 GGACAGCCCCAGTGAAGACACAGGAAATGGACCCCCACCTGGCCCCACCCAAACCTGGA 408
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
143 TGACAGCCTCTGTGAGACCCGAGGAAATGAACCCCTAATAAGGTTTCAACCCAGCCTGGA 202
```


QY 409 CACGTTTACCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTGACCAGAACCA 468
DB 203 CACATTTCACCTCTGAGGAGATGCTGCAGCAAAATGAAGGAACTCTGTTGAGAACCA 262
QY 469 GCTGAAGAAGCCATGAAGCTTAATATCAAGCCATGAAGGGAGATTGAGGAGCTTC 528
DB 263 GCTGAAGAAGCCATGAAGCTTAATCAATCAAGCCATGAAGGGAGATGAGGAGCTTC 322
QY 529 GGCTGGACAGAGAAAACAGAAAGGAAGACCGCAGTTTTTTTGAGATACAGAGCAAGAAC 588
DB 323 AGCTGGCAGAGAACGAGAGGAAACGACAGTTGTTTGACACACAGAACAAAGAAC 382
QY 589 AAAAGAGCGTCTAATGCGCTTGAATCATGAGATGAGAAATGAAGGAGCTTGGAAA 648
DB 383 TAGAAGCTCTGTAGTACCTTGAGTTATGAAATGAGAAATGAAAGGTGAGCTTGGAAA 442
QY 649 ACTAAAGGGGAAATCAGAAAGGTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAG 708
DB 443 ACTAAGAGAGAGTCAAGAAAGGCCAATTTGAAGGCTCCATATGTGGCTCCAGGCTCCCAT 502
QY 709 GG 710
DB 503 GG 504

RESULT 31

US-11-128-049-6348
; Sequence 6348, Application US/11128049
; Publication No. US2006010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; PRIOR FILING DATE: 2005-05-11
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 6348
; TYPE: DNA
; LENGTH: 505
; ORGANISM: Cricetulus griseus

Query Match 14.1%; Score 292.4; DB 16; Length 505;
Best Local Similarity 80.8%; Pred. No. 6.6e-70;
Matches 341; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 289 CTTTTCACAGGAACCTCTGCAATGTCCTCCATGCTCCATCAACCTCTCAGCTGCTCACTGAAAGGA 348
DB 83 CTAAGCCGACGGACCTGTTTACCATGTCCCATCAACCTCCGAGCTGCTGACTGAGAAGGG 142
QY 349 GGACAGCCCGAGTGAAGACAGAGAAATGGACCCCGCCCACTGCGCCCAACCAACCTGGA 408
DB 143 TGACAGCTCTGTGAGACCCCGAGAAATGAACCCCTTAATAGTTTCAACCCCGAGCTTGA 202
QY 409 CACGTTTACCCCGAGGAGCTGCTGCAGCAGATGAAGAGCTCTGTGACCAAGAACCA 468
DB 203 CACATTTCACCTCTGAGGAGATGCTGCAGCAAAATGAAGGAACTCTGTTGAGAACCA 262
QY 469 GCTGAAGAAGCCATGAAGCTTAATATCAAGCCATGAAGGGAGATTGAGGAGCTTC 528
DB 263 GCTGAAGAAGCCATGAAGCTTAATCAATCAAGCCATGAAGGGAGATGAGGAGCTTC 322

QY 529 GGCTGGACAGAGAAAACAGAAAGGAAGACCGCAGTTTTTTTGAGATACAGAGCAAGAAC 588
DB 323 AGCTGGCAGAGAACGAGAGGAAACGACAGTTGTTTGACACACAGAACAAAGAAC 382
QY 589 AAAAGAGCGTCTAATGCGCTTGAATCATGAGATGAGAAATGAAGGAGCTTGGAAA 648
DB 383 TAGAAGCTCTGTAGTACCTTGAGTTATGAAATGAGAAATGAAAGGTGAGCTTGGAAA 442
QY 649 ACTAAAGGGGAAATCAGAAAGGTCATCTGAGGACCCCACTGATGACTCCAGGCTTCCAG 708
DB 443 ACTAAGAGAGAGTCAAGAAAGGCCAATTTGAAGGCTCCATATGTGGCTCCAGGCTCCCAT 502
QY 709 GG 710
DB 503 GG 504

RESULT 32

US-09-960-352-11994
; Sequence 11994, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11994
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 51-LIB34-028-Q1-E1-E4
US-09-960-352-11994

Query Match 11.5%; Score 238.2; DB 3; Length 396;
Best Local Similarity 77.7%; Pred. No. 6.3e-55;
Matches 321; Conservative 0; Mismatches 73; Indels 19; Gaps 2;

QY 737 CTCAGGACCCAGGTGGTGAAGGTACAAGCAGAGAAAGCAGACCTGTTGGGCATCGTGCT 796
DB 1 CTGAGACAAAGGTGGCAGCCCTCCAGCTGAAAGGAGCATCTGCTGGGCATCGTGCT 60
QY 797 GAACTGCAGCTCAAGCTGAACTCCAGCGGCTCTCAGAGATTCTTTTGTGAAATTAGG 856
DB 61 GAATTGCAACTCAAGCTGAACTCAAGCGGCTCTCCGAGACTCTTTTGTGAAATCAGG 120
QY 857 ATGGCTGAAGGAGAACGAGAGGGTCAGTAAAGAAATCAAGCATAGTCTCTGGGCCACG 916
DB 121 ATGGGTGAAGGGGAGAGCAGCTGTGGGAGCACAAGAAATCAAGCAGATCTCTGGGCCACA 180
QY 917 AGAACAGTCTCCACTGCGACGGCATTGTCTAAATATAGGAGCAGATCTGCAGATGGGCG 976
DB 181 AGAACTGATTCTCGTTGACAC-----GAGCAATCTGCAGAGGTACC 222
QY 977 AGAATTAATCTCGAATGAGGAGTTAACTGTGAGCCAGCTCTCTGCTGCTCTAAGGAA 1036
DB 223 AGGAATTAATGTGGAGTTTGGGAAATTAATGTGAGCCAGCTCTGCTGATTTGTTCCAGGAA 282
QY 1037 GGAATCAGAGGTGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAGAAAGAGTTTCA 1096
DB 283 TGAACACAGAGGTGAGAGACTTGAAGTTGCTTCAAGGAGCCCAAGAAAGAAATTTCT 342
QY 1097 GATTTTGAAGAGAAACAAGTAACTCTTGAGA-TTGAACCCAGACAGAGG 1148
DB 343 GATTTTGAAGAGAAAGCCAGGATCATTCGAGACCGGAGACCAGACAGAGG 395

LOCATION: 3629
OTHER INFORMATION: single nucleotide polymorphism (SNP)
NAME/KEY: allele
LOCATION: 3882
OTHER INFORMATION: insertion of additional t residue
NAME/KEY: allele
LOCATION: 3988
OTHER INFORMATION: single nucleotide polymorphism (SNP)
NAME/KEY: allele
LOCATION: 4452
OTHER INFORMATION: single nucleotide polymorphism (SNP)
NAME/KEY: repeat region
LOCATION: 598..878
OTHER INFORMATION: repeat element
NAME/KEY: repeat region
LOCATION: 938..957
OTHER INFORMATION: Short repeat element
NAME/KEY: repeat region
LOCATION: 1002..1329
OTHER INFORMATION: ALU repeat element
NAME/KEY: repeat region
LOCATION: 2288..2587
OTHER INFORMATION: ALU repeat element
NAME/KEY: misc_feature
LOCATION: 5054
OTHER INFORMATION: putative transcription start site
US-10-091-281-2

Query Match 9.9%; Score 206; DB 7; Length 46951;
Best Local Similarity 100.0%; Pred. No. 7.9e-45;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 475 AGAAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTTCGGCCTG 534
DB 15116 AGAAGCCATGAAGCTAAATAATCAAGCCATGAAGGGAGATTGAGGAGCTTTTCGGCCTG 15175
QY 535 GACAGAAACAGAAAGAAAGCCAGCTTTTGTAGATACAGAGCAAGAAAGCAAAAGA 594
DB 15176 GACAGAAACAGAAAGAAAGCCAGCTTTTGTAGATACAGAGCAAGAAAGCAAAAGA 15235
QY 595 GCGTCTAATGGCTTGAGTCATGAGTAATGAGAAATGAGGAAGAGCTTGGAAAACTAAA 654
DB 15236 GCGTCTAATGGCTTGAGTCATGAGTAATGAGAAATGAGGAAGAGCTTGGAAAACTAAA 15295
QY 655 AGGGAATCAGAAAGTCTCTGAGG 680
DB 15296 AGGGAATCAGAAAGTCTCTGAGG 15321

RESULT 36
US-09-960-352-7573
Sequence 7573, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 7573
LENGTH: 283
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: unsure
LOCATION: (32), (218)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: 33-BOVMS1-004-Q1-E1-A2

US-09-960-352-7573
Query Match 9.8%; Score 202.6; DB 3; Length 283;
Best Local Similarity 82.0%; Pred. No. 3.9e-45;
Matches 232; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 1107 AGAAAAAAGTAACTCTTCTGAGATTGAAACCAGACAGAGGGGAGCAGAGAAAGAGA 1166
DB 1 AGAAAGCCCAAGGATCATTTCTGAGACTGAGACNCAGACAGAGGAGCAGACAGAAACAAGAGA 60
QY 1167 ATGATCAAGAGAAAGCCCGGAGACTGTTGGAAGCGAAGTGGAAAGCACTGAACCTCCAGG 1226
DB 61 AGAAGAGGAGAAAGACACCAAAACTATTGGAAGTGAAGTGAAGCACTGAACCTTCAGG 120
QY 1227 TGACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACAAACTCAGCGAAGCTGAGCTAA 1286
DB 121 TGACAAACCTGTTTAAAGGAGCTTCAAGAACTCAGCGAAGCTCAGTGAAGCTGAGCTAA 180
QY 1287 TGAAGAGAGACTTCAAGAAAGTGTGAGCCCTTGAAGGAAAAATTCGTGCAATTCAT 1346
DB 181 TGAAGAGAGACTTCAAGAAAAATGTGAGCCCTTGANAGGAAAAATCCGCAACCCAA 240
QY 1347 CAGAGTTGAATGAAGAAAGCAGAGCTTGTATTACTAACAAAA 1389
DB 241 CAAACTGAATGAAGAAAGCAGAGCTGGTTTATAATAAAACAAAA 283

RESULT 37
US-09-918-995-6208
Sequence 6208, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6208
LENGTH: 380
TYPE: DNA
ORGANISM: Homo sapiens
US-09-918-995-6208

Query Match 9.5%; Score 198; DB 3; Length 380;
Best Local Similarity 100.0%; Pred. No. 8.8e-44;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1880 AGAAAGTGACTCTGACCAGCAGGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGGAC 1939
DB 1 AGAAACAAGTGACTCTGACCAGCAGGCTTACCTTGTTCAAAGAGGAGCTGAGGACAGGGAC 60
QY 1940 TGGCGGCAACAGCGGAATATTCGATTCTCTGCCCAAGTGTGAGAGGTTCTGCCT 1999
DB 61 TGGCGGCAACAGCGGAATATTCGATTCTCTGCCCAAGTGTGAGAGGTTCTGCCT 120
QY 2000 GACATAGACAGCTTACAGATTACGCTGATGGATTGCATTTAAAGTGTGATGATCAC 2059
DB 121 GACATAGACAGCTTACAGATTACGCTGATGGATTGCATTTAAAGTGTGATGATCAC 180
QY 2060 CTCCCCAAAACTGTGGT 2077
DB 181 CTCCCCAAAACTGTGGT 198

RESULT 38
US-10-956-157-7613
Sequence 7613, Application US/10956157
Publication No. US20050118625A1

```

; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7613
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-7613

Query Match          9.3%; Score 194; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.5e-42;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1884 CAAAGTCTCTCAGCAGCAGGCTTACCTTGTTCAAAGAGGAGCTGAGGACGAGGACTGGC 1943
DB 1 CAAAGTCTCTCAGCAGCAGGCTTACCTTGTTCAAAGAGGAGCTGAGGACGAGGACTGGC 60

QY 1944 GGCACAGCGGGAATATTCGGATTTCCTCTGCCCCCAAGTGTGGAGAGGTTCTGCCTGACA 2003
DB 61 GGCACAGCGGGAATATTCGGATTTCCTCTGCCCCCAAGTGTGGAGAGGTTCTGCCTGACA 120

QY 2004 TAGACACGTTACAGATTACGTTACGTTAGTGATGATCATATTAAGTGTGATATCATCTCC 2063
DB 121 TAGACACGTTACAGATTACGTTACGTTAGTGATGATCATATTAAGTGTGATATCATCTCC 180

QY 2064 CCAAAACTGTTGGT 2077
DB 181 CCAAAACTGTTGGT 194

RESULT 39
US-10-627-757-2
; Sequence 2, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-2

Query Match          9.2%; Score 190.4; DB 8; Length 1166;
Best Local Similarity 99.5%; Pred. No. 2.1e-41;
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 285 GTGACTTTCCACAGGAATCTTGCATGTCCCATCAACTCTCAGTCGCTCACTGAAA 344
DB 475 GTGACTTTTCCACAGGAATCTTGCATGTCCCATCAACTCTCAGTCGCTCACTGAAA 534

QY 345 AGGAGGACAGCCCGAGTGAAGCAGAGAAATGGACCCCGCCACCTGGCCCAACCCAAACC 404
DB 535 AGGAGGACAGCCCGAGTGAAGCAGAGAAATGGACCCCGCCACCTGGCCCAACCCAAACC 594

QY 405 TGGACACGTTTACCCCGAGGAGCTGCTCAGCAGATGAAGAGCTCTTGACCAAGAAC 464
DB 405 TGGACACGTTTACCCCGAGGAGCTGCTCAGCAGATGAAGAGCTCTTGACCAAGAAC 464
```

```

DB 595 TGGACACGTTTACCCCGAGGAGCTGCTCAGCAGATGAAGAGCTCTTGACCAAGAAC 654
QY 465 ACCAGCTGAAG 476
DB 655 ACCAGCTGAAG 666

RESULT 40
US-10-627-757-4
; Sequence 4, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-4

Query Match          9.0%; Score 186; DB 8; Length 1183;
Best Local Similarity 100.0%; Pred. No. 3.6e-40;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGACCCCACTGATGATCTCCAGGCTTCCAGGGCCGAGGACGAGGAGGACGAGC 737
DB 499 AGGACCCCACTGATGATCTCCAGGCTTCCAGGGCCGAGGACGAGGAGGAGGACGAGC 558

QY 738 TCAGGACCCAGGTGTGAGGCTTACAGCAGAGAGGACGAGCTGTGGGCATCTGCTCTG 797
DB 559 TCAGGACCCAGGTGTGAGGCTTACAGCAGAGAGGACGAGCTGTGGGCATCTGCTCTG 618

QY 798 AACTGCAGCTCAAGCTGAACCTCCAGCGGCTCCTCAGAGATTCTTTGTTGAAATTAGGA 857
DB 619 AACTGCAGCTCAAGCTGAACCTCCAGCGGCTCCTCAGAGATTCTTTGTTGAAATTAGGA 678

QY 858 TGGCTG 863
DB 679 TGGCTG 684

RESULT 41
US-10-627-757-11
; Sequence 11, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1159
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-11

Query Match          7.8%; Score 162.4; DB 8; Length 1159;
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-6

Query Match 7.5%; Score 155.8; DB 8; Length 1153;
Best Local Similarity 88.5%; Pred. No. 8.4e-32;
Matches 169; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 934 CACGGCATTGCTTAATATATAGGAGCAGATCTGCAGATGGGGCCCAAGAAATTTACTTCAACA 993
    |||
Db 498 CAGGGCATTGCTTAATATATAGGAGCAGATCTGCAGATGGGGCCCAAGAAATTTACTTCAACA 557

QY 994 TGAGGAGTTAACTGTGAGCCAGCTCTCTGTGTCCTTAAGGGAAGGGAATCAGAAAGTGA 1053
    |||
Db 558 TGAGGAGTTAACTGTGAGCCAGCTCTCTGTGTCCTTAAGGGAAGGGAATCAGAAAGTGA 617

QY 1054 GAGACTTGAAGTTGCACTCAAGGAGGCCCAAGAAAGTTTCAGATTTTGAAGAAAC 1113
    |||
Db 618 GAGACTTGAAGTTGCACTCAAGGAGGCCCAAGAAAGTTTCAGATTTTGAAGAAAT 677

QY 1114 AAGTAATCGTT 1124
    |||
Db 678 ATGTGTTTTTT 688

RESULT 44
US-10-627-757-9
; Sequence 9, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-9

Query Match 7.3%; Score 152.4; DB 8; Length 1150;
Best Local Similarity 99.4%; Pred. No. 7.3e-31;
Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1307 AAGTGTGAGCCCTTGAAAGGAAATTTCTGCAATTCATCAGATTTGAATGAAAGCAA 1366
    |||
Db 499 AGGTGTGAGCCCTTGAAAGGAAATTTCTGCAATTCATCAGATTTGAATGAAAGCAA 558

QY 1367 GAGCTTGTGTTTACTAATAACAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATC 1426
    |||
Db 559 GAGCTTGTGTTTACTAATAACAAAAGTTAGAGCTACAGTGGAAAGCATGCTATCAGAAATC 618

QY 1427 AAAATGGAACAGCTTAAAAACAGAGGATGAAAGT 1460
    |||
Db 619 AAAATGGAACAGCTTAAAAACAGAGGATGAAAGT 652

RESULT 45
US-10-627-757-12
; Sequence 12, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-12

Query Match 7.6%; Score 157; DB 8; Length 1122;
Best Local Similarity 100.0%; Pred. No. 3.8e-32;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1921 AGGAGCTGAGCAGAGGAGCTGGCGCAACAGCGGAATATTCGATTCATTCCTGCCCCAA 1980
    |||
Db 499 AGGAGCTGAGCAGAGGAGCTGGCGCAACAGCGGAATATTCGATTCATTCCTGCCCCAA 558

QY 1981 GTGTGGAGAGTCTGCTGACATAGACACGTTACAGATTCACGTGATGATTCATCAT 2040
    |||
Db 559 GTGTGGAGAGTCTGCTGACATAGACACGTTACAGATTCACGTGATGATTCATCAT 618

QY 2041 TTAAGTGTGATGATATCACCTCCCAAACTGTTGGT 2077
    |||
Db 619 TTAAGTGTGATGATATCACCTCCCAAACTGTTGGT 655

RESULT 43
US-10-627-757-6
; Sequence 6, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1153
```

```
/ FILE REFERENCE: Q76319
/ CURRENT APPLICATION NUMBER: US/10/627,757
/ CURRENT FILING DATE: 2003-07-28
/ PRIOR APPLICATION NUMBER: JP P2002-226612
/ PRIOR FILING DATE: 2002-08-02
/ NUMBER OF SEQ ID NOS: 40
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 12
/ LENGTH: 1131
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-627-757-12

Query Match          6.4%; Score 135.4; DB 8; Length 1131;
Best Local Similarity 99.3%; Pred. No. 3.8e-26;
Matches 136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1709 CAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCGAGAGAGAAATTCAT 1768
Db 498 CAGATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCGAGAGAGAAATTCAT 557
QY 1769 GAGGAAAAGAGCAACTGCGATTGCGAGCTGGCAGTTCTGCTGAAGAGAGATGATGCTTTC 1828
Db 558 GAGGAAAAGAGCAACTGCGATTGCGAGCTGGCAGTTCTGCTGAAGAGAGATGATGCTTTC 617
QY 1829 GAAGACGAGGAGCGCA 1845
Db 618 GAAGACGAGGAGCGTA 634

RESULT 46
US-10-750-185-4641
/ Sequence 4641, Application US/10750185
/ Publication No. US20050260603A1
/ GENERAL INFORMATION:
/ APPLICANT: MMI GENOMICS, INC.
/ APPLICANT: DENISE, Sue K.
/ APPLICANT: KERR, Richard
/ APPLICANT: ROSENFELD, David
/ APPLICANT: HOLM, Tom
/ APPLICANT: BATES, Stephen
/ APPLICANT: FANTIN, Dennis
/ TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
/ FILE REFERENCE: MM1100-2
/ CURRENT APPLICATION NUMBER: US/10/750,185
/ CURRENT FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437,482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 4641
/ LENGTH: 600
/ TYPE: DNA
/ ORGANISM: Bovine MMBT07354
US-10-750-185-4641

Query Match          6.4%; Score 133.2; DB 10; Length 600;
Best Local Similarity 80.4%; Pred. No. 1.1e-25;
Matches 156; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 678 AGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAGCGGAGGAGGAAAGGACCAGC 737
Db 362 AGGACATCATTTGGGACCCCAAGTCCCAAGCGCAGAGTGGAAACAAGAGTGGAAACAGC 421
QY 738 TCAGGACCCAGGTGGTGAGGCTTACAAGCAGAGAGGAGAGCTGTTGGGCATCGTGCTG 797
Db 422 TGAAGACCCAGGTGGTGAGGCTTCCAAGCTCCAAAGCGAGATCTGCTGGGCATCGTGCTG 481
QY 798 AACTGCAGCTCAAGCTGAAGTCCAGCGGCTCCTCAGAAGATTCCTTTGTTGAAATTAGGA 857
Db 482 AATTGCAACTCAAGCTGAAGTCCAGCGGCTCCTCCGAGAGACTCCTTTGTTGAAATCAGGA 541
QY 858 TGGCTGAAGAGGAA 871
Db 542 TGGCTGTAAGTGTA 555

RESULT 48
US-10-627-757-10
/ Sequence 10, Application US/10627757
/ Publication No. US20040091914A1
/ GENERAL INFORMATION:
/ APPLICANT: KOUCHI YASUHIRO
/ APPLICANT: MASASGO AKINORI
/ APPLICANT: TAKAHATI TAKAYUKI
/ TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
/ FILE REFERENCE: Q76319
/ CURRENT APPLICATION NUMBER: US/10/627,757
/ CURRENT FILING DATE: 2003-07-28
/ PRIOR APPLICATION NUMBER: JP P2002-226612
/ PRIOR FILING DATE: 2002-08-02
/ NUMBER OF SEQ ID NOS: 40
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 10
/ LENGTH: 1094
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-627-757-10
```

```
Db 542 TGGCTGTAAGTGTA 555

RESULT 47
US-10-750-623-4641
/ Sequence 4641, Application US/10750623
/ Publication No. US20050287531A1
/ GENERAL INFORMATION:
/ APPLICANT: MMI GENOMICS, INC.
/ APPLICANT: DENISE, Sue K.
/ APPLICANT: KERR, Richard
/ APPLICANT: ROSENFELD, David
/ APPLICANT: HOLM, Tom
/ APPLICANT: BATES, Stephen
/ APPLICANT: FANTIN, Dennis
/ TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
/ FILE REFERENCE: MM1100-1
/ CURRENT APPLICATION NUMBER: US/10/750,623
/ CURRENT FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437,482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 4641
/ LENGTH: 600
/ TYPE: DNA
/ ORGANISM: Bovine MMBT07354
US-10-750-623-4641

Query Match          6.4%; Score 133.2; DB 10; Length 600;
Best Local Similarity 80.4%; Pred. No. 1.1e-25;
Matches 156; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 678 AGGACCCCACTGATGACTCCAGGCTTCCAGGGCCGAGCGGAGGAGGAAAGGACCAGC 737
Db 362 AGGACATCATTTGGGACCCCAAGTCCCAAGCGCAGAGTGGAAACAAGAGTGGAAACAGC 421
QY 738 TCAGGACCCAGGTGGTGAGGCTTACAAGCAGAGAGGAGAGCTGTTGGGCATCGTGCTG 797
Db 422 TGAAGACCCAGGTGGTGAGGCTTCCAAGCTCCAAAGCGAGATCTGCTGGGCATCGTGCTG 481
QY 798 AACTGCAGCTCAAGCTGAAGTCCAGCGGCTCCTCAGAAGATTCCTTTGTTGAAATTAGGA 857
Db 482 AATTGCAACTCAAGCTGAAGTCCAGCGGCTCCTCCGAGAGACTCCTTTGTTGAAATCAGGA 541
QY 858 TGGCTGAAGAGGAA 871
Db 542 TGGCTGTAAGTGTA 555

RESULT 48
US-10-627-757-10
/ Sequence 10, Application US/10627757
/ Publication No. US20040091914A1
/ GENERAL INFORMATION:
/ APPLICANT: KOUCHI YASUHIRO
/ APPLICANT: MASASGO AKINORI
/ APPLICANT: TAKAHATI TAKAYUKI
/ TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
/ FILE REFERENCE: Q76319
/ CURRENT APPLICATION NUMBER: US/10/627,757
/ CURRENT FILING DATE: 2003-07-28
/ PRIOR APPLICATION NUMBER: JP P2002-226612
/ PRIOR FILING DATE: 2002-08-02
/ NUMBER OF SEQ ID NOS: 40
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 10
/ LENGTH: 1094
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-627-757-10
```


Query Match 6.0%; Score 124; DB 8; Length 1094;
Best Local Similarity 100.0%; Pred. No. 5.4e-23;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1337 GCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTATATCTAACAATAAGTTAGAG 1396
Db 1 GCAATTCATCAGAGTTGAATGAAAGCAAGAGCTTGTATATCTAACAATAAGTTAGAG 60
QY 1397 CTACAAGTGAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAACACAGAGCATGAA 1456
Db 61 CTACAAGTGAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAACACAGAGCATGAA 120
QY 1457 AAGT 1460
Db 121 AAGT 124

RESULT 49
US-09-960-352-14197
; Sequence 14197, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 14197
; LENGTH: 170
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (18),(141)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 61-BOVMS1-018-Q1-E1-H2
US-09-960-352-14197

Query Match 5.8%; Score 120; DB 3; Length 170;
Best Local Similarity 85.3%; Pred. No. 2.4e-22;
Matches 145; Conservative 0; Mismatches 22; Indels 3; Gaps 1;
QY 1727 TCTGATTTTCATGCTGAAAGAGCGAGAGAGAGAAATTCATGAGGAAAGAGGCAACTG 1786
Db 1 TCTGATTTTCATGCTGAAAGAGCGAGAGAGAGAAATTCATGAGGAAAGAGGCAACTG 60
QY 1787 GCATTGCGAGTGGCAGTTCTGCTCAAGAGAAATGCTTTTCGAGAGCGAGGCGAG---G 1843
Db 61 GCATTGCGAGTGGCAGTTCTGCTCAAGAGAAATGCTTTTCGAGAGCGAGGCGAG 120
QY 1844 CAGTCCTTGATGGAGATGAGAGTCGTCATGGGGCGAGAGCAAGTGAATC 1893
Db 121 CAGACCTTGATGGAAATGCANAGCCGTACGGGGCGAGAGCAAGTGAATC 170

RESULT 50
US-10-627-757-8
; Sequence 8, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASCO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: 076319
; CURRENT APPLICATION NUMBER: US/10/627,757
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612

; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-8
Query Match 5.6%; Score 116.4; DB 8; Length 1116;
Best Local Similarity 99.2%; Pred. No. 7e-21;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1193 GTTGGAAAGCGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAA 1252
Db 501 GTTGGAAAGCGAAGTGAAGCACTGAACCTCCAGGTGACATCTCTGTTTAAAGGAGCTTCAA 560
QY 1253 GAGGCTCATACAAAACTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAAGT 1310
Db 561 GAGGCTCATACAAAACTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAAGT 618
Search completed: May 30, 2006, 00:41:42
Job time : 2644 secs

THIS PAGE BLANK (USPTO)


```
199 32.2 1.6 1658 7 US-11-101-316-59
100 32.2 1.6 1755 7 US-11-217-529-81374
101 32.2 1.6 2403 7 US-11-217-529-78558
102 32.2 1.6 3639 7 US-11-217-529-80622
103 32.2 1.6 3693 7 US-11-217-529-77316
104 32.2 1.6 4359 7 US-11-314-018-3
105 32.2 1.6 8370 7 US-11-263-326-90
106 32.2 1.6 8377 7 US-11-263-326-95
107 32.2 1.6 8380 7 US-11-263-326-89
108 32.2 1.6 8381 7 US-11-263-326-1
109 32.2 1.6 8381 7 US-11-263-326-3
110 32.2 1.6 8381 7 US-11-263-326-5
111 32.2 1.6 8381 7 US-11-263-326-7
112 32.2 1.6 8381 7 US-11-263-326-8
113 32.2 1.6 8381 7 US-11-263-326-9
114 32.2 1.6 8381 7 US-11-263-326-11
115 32.2 1.6 8381 7 US-11-263-326-91
116 32.2 1.6 8381 7 US-11-263-326-92
117 32.2 1.6 8381 7 US-11-263-326-93
118 32.2 1.6 8381 7 US-11-263-326-96
119 32.2 1.6 8381 7 US-11-263-326-97
120 32.2 1.6 8381 7 US-11-263-326-108
121 32.2 1.6 8381 7 US-11-263-326-136
122 32.2 1.6 8381 7 US-11-263-326-137
123 32.2 1.6 8381 7 US-11-263-326-138
124 32.2 1.6 8381 7 US-11-263-326-139
125 32.2 1.6 8381 7 US-11-263-326-140
126 32.2 1.6 8381 7 US-11-263-326-141
127 32.2 1.6 8381 7 US-11-263-326-142
128 32.2 1.6 8381 7 US-11-263-326-143
129 32.2 1.6 8381 7 US-11-263-326-144
130 32.2 1.6 8388 7 US-11-263-326-94
131 32 1.5 252 7 US-11-217-529-174207
132 32 1.5 759 7 US-11-217-529-79408
133 32 1.5 1172 6 US-10-953-349-23583
134 32 1.5 1937 1 US-09-949-925-61
135 32 1.5 1986 7 US-11-217-529-2475
136 32 1.5 2171 7 US-11-293-697-1265
137 32 1.5 2269 7 US-11-293-697-2243
138 32 1.5 3356 7 US-11-293-697-457
139 31.8 1.5 984 7 US-11-217-529-1241
140 31.8 1.5 1106 7 US-11-253-300-2
141 31.8 1.5 1605 7 US-11-217-529-78602
142 31.8 1.5 2040 6 US-10-953-349-37996
143 31.8 1.5 2193 6 US-10-953-349-11158
144 31.8 1.5 2508 7 US-11-217-529-191162
145 31.8 1.5 54550 7 US-11-318-813-42
146 31.6 1.5 321 7 US-11-217-529-2470
147 31.6 1.5 1099 6 US-10-953-349-18047
148 31.6 1.5 1147 6 US-10-953-349-6314
149 31.6 1.5 1625 6 US-10-953-349-17058
150 31.6 1.5 1903 6 US-10-505-928-806
```

ALIGNMENTS

RESULT 1

```
US-11-217-529-79814
; Sequence 79814, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182

Query Match 2.1%; Score 42.6; DB 7; Length 1803;
Best Local Similarity 47.5%; Pred. No. 0.007;
Matches 126; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 1435 ACAGCTAAAAACAGAGGATGAAAGTCCAAATTAACCTGTGTGTACAGACACACAA 1494
DB 30 ACAGCTGAAAGGAGGCCACGAAATAGTATCAAGGCTAGAAAGTACAGACAGATAA 89
QY 1495 GCTTCTTCAAGACATAATAATGCTTGAACCAATTTGAGGAACCTAACAGAAAAGATC 1554
```

```
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79814
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-79814

Query Match 2.1%; Score 44; DB 7; Length 1803;
Best Local Similarity 46.2%; Pred. No. 0.0078;
Matches 146; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 1487 CACACACAGCTTCTTCAAGAACATAATAATGCTTGAACCAATTTGAGGAACCTAACACAA 1546
DB 1042 CAAAAGACATTCAGTTTCAAAAGGAAATAGATGAAAAGATTAGAAAGGAAAGAA 1101
QY 1547 AAAGAGTCAGAAAAAGTGGACAGGCGCAGTGTCTGAAGGAACCTGAGTGAAGAACTTGAACCTG 1606
DB 1102 ATAAAGCAGGAGGAGAGGAAAAAGGCAATGGAATTTGGAACCAACCAACGCACTCGAAGAA 1161
QY 1607 GCAGAGAGGCTCTGGCTTCCAAACAGCTGCAAAATGGATGAAATGAAGCAACCACTTGGC 1666
DB 1162 GAGGAACGAAAAACGTCAAATGGAATCTGAGGCTAAAAAGCAATGGAGTTGAAACGCTCAA 1221
QY 1667 AAGCAGGAAGAGGAGGACCTGGAACCATGACCATCTCTCAGGGCTCAGATGGAAGTTTACTGT 1726
DB 1222 CGGCAATTCGAAAGAGAGCAACGCTTAAAAAGAGAGAGAACTATTGGAATATACAGAGA 1281
QY 1727 TCTGATTTTCATGCTGAAAGAGCGAGAGAGAGAAAATTCATGAGGAAAAAGGAGCAACTG 1786
DB 1282 AAACAAAGAGAAACAGAAACAGCTGAGAGATTAAAAAAGAGCAAGGCGCTTGTCT 1341
QY 1787 GCATTCGAGCTGGCAG 1802
DB 1342 AAAAAGGAGGAGGAAG 1357
```

RESULT 2

```
US-11-217-529-166820
; Sequence 166820, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 166820
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-166820
```

```
Query Match 2.1%; Score 42.6; DB 7; Length 339;
Best Local Similarity 47.5%; Pred. No. 0.007;
Matches 126; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 1435 ACAGCTAAAAACAGAGGATGAAAGTCCAAATTAACCTGTGTGTACAGACACACAA 1494
DB 30 ACAGCTGAAAGGAGGCCACGAAATAGTATCAAGGCTAGAAAGTACAGACAGATAA 89
QY 1495 GCTTCTTCAAGACATAATAATGCTTGAACCAATTTGAGGAACCTAACAGAAAAGATC 1554
```

Db	90	GTTGAACGACCAAGACTGATGACGCCAGGAATCGACTCATACAAATTCAAAAAGA	149
Qy	1555	AGAAAAAGTGGACACGGCCAGTGTGTAAGGAACCTGAGTGA AAAA CTGGAACCTGGCAGAGAA	1614
Db	150	CAAGGAATTGAAGGAGTTTGAACAAAAGAAATGCCGGTGGTCTTGGTGAATTTGGA AAAA AGAA	209
Qy	1615	GGCTCTGGCTTCCAAACAGCTGC AAA TGGATGAAATGCAAGCAAAACCAATTTGCCAAGCAGGGA	1674
Db	210	AGCAGAGGCTGCTGTGCAAGGTGAATTAGCTGAGATTAAAGAAAATTCGACAGAAAGAAAA	269
Qy	1675	AGAGGACCTGGAACCAATGACCATC	1699
Db	270	GGATGACGTTGTCAAAAATTTTGTATC	294

```

RESULT 3
US-11-217-529-190968
; Sequence 190968, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 190968
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-190968

```

RESULT 4
US-11-217-529-6230
; Sequence 6230, Application US/11217529
; Publication No. US2006099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIRO

```

; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 8-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6230
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-6230

```

```

RESULT 5
US-11-217-529-77435/c
; Sequence 77435, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77435
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (152)..(192)
; OTHER INFORMATION: a, c, g, t, unknown, or other
US-11-217-529-77435

```

	Query Match	1.9%	Score 40.4;	DB 7;	Length 615;
	Best Local Similarity	40.2%;	Pred. No. 0.043;		
	Matches 149;	Conservative 0;	Mismatches 222;	Indels 0;	Gaps 0;
QY	1039	GAATCAGAGGTGGAGAGACTTGAAGTTGCACTCAGGAGGCCAAGAAAGAGTTTCAGA			1098
Db	458	GAGTTTATAGAGGATAGTAGTAGAGAGAAATAGAAAGAGTAGAAGAGAGAGAAAT			399
QY	1099	TTTTTGAAGAAACCAAGTAATCGTCTCGAGATTGAAACCAGACAGAGGGGAGCACAGA			1158


```
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-5260

Query Match      1.9%; Score 39.6; DB 7; Length 2832;
Best Local Similarity 49.1%; Pred. No. 0.18; 109; Indels 0; Gaps 0;
Matches 105; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 1281 AGCTAATGAGAGAGAGCTTCAAGAAAGTGTCTAGGCCCTTGAAGAGAAATTTCTCAA 1340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1325 AGGAATAGACAGAGAGCTTCAAGAAACATTTAAACGATGAGGATAGTTACTCGAGTAGCA 1384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1341 TTCCATCAGAGTTGAATGAAAGCAAGAGCTTGTATATCTAACAAGAGTTAGAGTAC 1400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 TAGCAAGACCATAGATAAAATTTCTGAAATAGTTGAAATATTAATGAAGAGTTGGACCAAC 1444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1401 AAGTGGAAAGCATCTCTAGAAATCAAAATGGAACAGGCTAAACACAGAGATGAAAAAGT 1460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1445 ATGTGGATGAATCCAGTATGCAACGCGCTGAACAAAAATGAAAAAGATGAGTCTG 1504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1461 CCAATTAAGTGTCTACAGATGACACACAA 1494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1505 CCGTATTCAATGAATGAGTGCCTCCGCAACGA 1538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-11-217-529-79917
; Sequence 79917, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 79917
; LENGTH: 2007
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-79917

Query Match      1.9%; Score 39.4; DB 7; Length 2007;
Best Local Similarity 45.4%; Pred. No. 0.17;
Matches 142; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 1248 TTCAGAGGCTCATACAAACTCAGCAAGCTGAGCTAATGAAGAGAGACTTCAAGAA 1307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 TTAACATAAACATACATCAATAAGAGAACTCTGGAGAACGGTGAAGCTCGTATGAAA 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1308 AGTGTAGGCCCTTGAAGAGAAATTTCTGCAATCCATCAGAGTTGAAATGAAAGCAAG 1367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 TGTTAATGAATTTGTTAAGAAACCTTCCGAAGAAATAGACAGTTTGAATGAGGATGTCG 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1368 AGCTTGTTTATCTAACAAAGTTAGCTACAAAGTGAAGCATGCTATCAGAAATCA 1427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 AGGAATCAAACTTCAGAAAGAAAGAACTACACAAAGAGGCACAACTTCCAAGTTGG 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1428 AAATGGACAGGCTAAAACAGAGATGAAAGTCCAAATTAATGTGCTACAGATGACAC 1487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 AAAATAAGGAGACGGTCTTTTAAATGAAATTTCTCGTTTGAAGAACTTCACTGTCGATA 964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1488 ACAACAGCTTCTTCAAGACATAAATGCAATGAAACAAATTTAGGAGACTAACAGAA 1547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 AAGTAGAAATTTGAATGATACTACTGAAATAATCAAGGCTCTGGAGTCAAGAAATTCGCA 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Query Match      1.9%; Score 39.4; DB 6; Length 394191;
Best Local Similarity 50.8%; Pred. No. 4.2;
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1384 CAAAAGTTAGAGCTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAA 1443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163910 CAGAAAGAGAAATATAGTTCAGAAATTTTGATCTATATAAGAAAGAGAGAAAGAAA 163851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1444 ACAGAGGATGAAAGTCCAAATTAAGTGTGCTACAGATGACACACACAGCTTTCTCA 1503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163850 TAAAGAGATAAAACCTTTAAATTTTTCCTACTTTTAAGATGATCAGAAAGCTTTGTTCA 163791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1504 AGAACATAATAATGCAATTCGAAACCAATTCAGGAACCTAACAGAAAGAGTCAGAAAAAGT 1563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163790 AAATACACAACTGATTTTGATACCTATAGTTTACGACACAGTGAATGAATGACAAAA 163731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1564 GGACA 1568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163730 TGATA 163726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-11-217-529-82102/c
; Sequence 82102, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 82102
; LENGTH: 1212
```

```
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-82102

Query Match          1.9%; Score 39; DB 7; Length 1212;
Best Local Similarity 47.4%; Pred. No. 0.16;
Matches 117; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 965 GCAGATGGGGCCCAAGAAATTAACCTCGAACATGAGGAGTAACTGTGAGCCAGCTCTCTGCTG 1024
Db 1196 GAAGAGAAAGCCACTGTTTCTTGGATGAAGAGGAGGATGCTTTTCGAGTGAACATAATG 1137

QY 1025 TGCCTAAGGGAAGGGGAATCAGAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAA 1084
Db 1136 GAACATAGATGTGTGCTGATGAGGAAGCAAGTGTTTTAGATGAAGTACGGTTCGAAGAA 1077

QY 1085 GAAAGAGTTTCAGATTTTGAAGAAGAAAACAAGTAATCGTTCTGAGATTGAAACCCAGACA 1144
Db 1076 GATGGAGATGTGATGAAGAGGAAGATACAGAAGACGAAGAAGATACAGTGGGAAGAACCC 1017

QY 1145 GAGGGGACACAGAGAAAGAAATGATGAAGAGAAAGGCCCGGAGACTGTTTGGAGCGAA 1204
Db 1016 GAGGAAGATACAGTGGGAAGACCCGAGGAAGATACAGTGGGAAGAAACCGAGGAAGATACA 957

QY 1205 GTGGAAG 1211
Db 956 GTGGAAG 950

RESULT 12
US-11-217-529-78674
; Sequence 78674, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 78674
; LENGTH: 2466
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78674

Query Match          1.9%; Score 38.8; DB 7; Length 2466;
Best Local Similarity 45.4%; Pred. No. 0.28;
Matches 139; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 1281 AGCTAATGAAGAAGAGACTTCAAGAAAAAGTGTGAGGCCCTTGAAAGGAAAAATTCGTCAA 1340
Db 1610 AGGTAGTACAGGATAAATTAATTAATAATGAGGAGACGCTAAAACTAAAGAAGAGCTGAAA 1669

QY 1341 TTCCATCAGATTGTAATGAAGACAGAGCTGTTTATCTACTACAAAAGATTAGAGCTAC 1400
Db 1670 TTGAATTCATTAATAGTGAATGGAGCGTGAAGAAACAGATCACTCAAAAAGATGATG 1729

QY 1401 AAGTGGAAAGCATGCTATCAGAAATCAAAATGAAACAGGCTAAAAACAGAGGATGAAAAAGT 1460
Db 1730 AATTTAAGATGTGGCAAGCAAAATACGAAACTGTGCAAGACGAAGCTAAGATACGCAATG 1789

QY 1461 CCMAATTAACTGTGCTACAGATGACACACACAGCTTCTTCAAGAACATATAATGCAAT 1520
Db 1790 CCGAAGTTACAGAGTTGAACCGGAGATATTGAAGATCTAAAGGAATCAAAATTTGCATTTGG 1849
```

```
QY 1521 TGAAAAACAATTGAGGAACATAACAAGAAAAAGAGTCAGAAAAAGTGACAGGCGAGTCTGCA 1580
Db 1850 AAGAAAAACAATTACGGAACCTAGAAAAACAAGTTCATAAACTAGAAAAATGAGTGTGAACGG 1909

QY 1581 AGGAAC 1586
Db 1910 AGAAAC 1915

RESULT 13
US-11-217-529-174162/c
; Sequence 174162, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 174162
; LENGTH: 768
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-174162

Query Match          1.8%; Score 38.2; DB 7; Length 768;
Best Local Similarity 46.7%; Pred. No. 0.21;
Matches 121; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 965 GCAGATGGGGCCCAAGAATTACTTCGAACATGAGGAGTTAACTGTGAGCCAGCTCTCTGCTG 1024
Db 422 GAAGAGGAAGCCACTGTTTCTTGTGATGAAGAGGAGGATTCGTTTTCGAGTGAACATAATG 363

QY 1025 TGCCTAAGGGAAGGGAATCAGAAGGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAA 1084
Db 362 GNACTAGATGTGTGGCTGATGAGGAAGCAAGTGTTTTAGATGAAGTAGCGGTCCGAAGAA 303

QY 1085 GAAAGAGTTTCAGATTTTGAAGAAAAACAAGTAATCGTTCTGAGATTGAAACCCAGACA 1144
Db 302 GATGGAGATGTGGATGAAGAGGAAGATACAGAAGACGAAGAAGATACAGTGGGAAGAACCC 243

QY 1145 GAGGGGACACAGAGAAAGAGAAATGATGAAGAGAAAGCCCGGAGACTGTTTGGAGCGGAA 1204
Db 242 GAGGAAGACACAGTGGAAAGACACCGTGGGAAGACGCCGGAAGGAGATACAGTAGAAGAGGAT 183

QY 1205 GTGGAAGCACTGAACCTCC 1223
Db 182 GTAGAGCTGCTGGAATATAC 164

RESULT 14
US-10-953-349-17644/c
; Sequence 17644, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patent in version 3.3
```

; SEQ ID NO 17644
; LENGTH: 1285
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-17644

Query Match 1.8%; Score 38.2; DB 6; Length 1285;
Best Local Similarity 51.5%; Pred. No. 0.28;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 1253 GAGGCTCATCAAACTCAGCGAGCTGAGCTAATGAGAGAGACTTCAAGAAAAGTGT 1312
DB 1103 GAGCACCAAAATGCTTTTCAGCTAAGCCAGGTACCATCAACTTTGAAATGGAATTT 1044
QY 1313 CAGGCCCTTGAAGAGAAAATTTCTGCAATTCATCAGAGTTGAATGAAAAGCAAGAGCTT 1372
DB 1043 GCAGCATCTTCTGACAACTTCTCTCAGTTCCATGTGAATCAGATGAAGAGCTAGAGCAT 984
QY 1373 GTTATACTAACAAAAGTTAGAGCTACAAGTGAAGATGCTATCAGAA 1423
DB 983 CTTCCCAATAACATACCTCTGTCAACCCCAAGTAGATTGAATTTCTATGAAA 933

RESULT 15

US-11-217-529-81089
; Sequence 81089, Application US/11217529
; Publication No. US20060099612A1

; GENERAL INFORMATION:

; APPLICANT: SUNTORY LIMITED

; APPLICANT: NAKAO, YOSHIHIRO

; APPLICANT: NAKAMURA, NORIHISA

; APPLICANT: KODAMA, YUKIKO

; APPLICANT: FUJIMURA, TOMOKO

; APPLICANT: ASHIKARI, TOSHIHIKO

; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS

; FILE REFERENCE: S-38-285

; CURRENT APPLICATION NUMBER: US/11/217,529

; PRIOR FILING DATE: 2005-09-02

; PRIOR APPLICATION NUMBER: US 10/932,182

; NUMBER OF SEQ ID NOS: 2004-09-02

; SOFTWARE: Patent in version 3.3

; SEQ ID NO 81089

; LENGTH: 2796

; TYPE: DNA

; ORGANISM: Saccharomyces pastorianus

US-11-217-529-81089

Query Match 1.8%; Score 38; DB 7; Length 2796;
Best Local Similarity 44.6%; Pred. No. 0.52;
Matches 149; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 1096 AGATTTTGAAGAAAACAAAGTATCGTTCTGAGATTGAACCCAGACAGAGGGGAGCAC 1155
DB 960 AGAACTAGAGAGAGAAAAGAAAAGAAAAGAGAGAGAAAAGAAAAGCAAAAAAAG 1019
QY 1156 AGAAGAAAGAGATGATGAAGAGAAAGCCCGGAGACTGTTGGAAGCGAGTGAAGCACT 1215
DB 1020 AGAAAAGAGAGAAAAGAAAAGAGAGACTGCAACAGTTAGCTAAAGAGAGAAAAGCAAA 1079
QY 1216 GAACCTCCAGTGACATCTCTGTTTAAAGAGCTTCAAGAGGCTCATACAAAACCTCAGCGA 1275
DB 1080 AAGGGAAGAGAGAAAAGAAAAGGCTTAAAGAGGAATTAGAAGAGCGTGAATGAGAAGCG 1139
QY 1276 AGCTGAGCTAATGAAGAGAGACTTCAAGAAAAGTGTCAAGGCCCTTGAAGGAAAATTC 1335
DB 1140 AGAGGCTCAAGGAAAAGAGGTTGAAGAGCCAAACGGAAGAAAAGCAGGAGAGAAAAGCG 1199
QY 1336 TGCAATTTCCATCAGAGTTGAATGAAAAGCAAGAGCTGTTTATACTAACAAAAGTTAGA 1395
DB 1200 TAGATTGGAAGAGCAGCGGAGAGAGAAATGCAAGAAAACAAACGGAAGAAAAGAGA 1259
QY 1396 GCTACAAGTGGAAAGCATGCTATCAGAAAATCAAA 1429

DB 1260 GGAACCTAAAACGCAAGAGGGAAGAAAGAGAAA 1293

RESULT 16

US-10-953-349-26062
; Sequence 26062, Application US/10953349
; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: ENCODED THERBY

; FILE REFERENCE: 2750-1579PUS2

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: Patent in version 3.3

; SEQ ID NO 26062

; LENGTH: 883

; TYPE: DNA

; ORGANISM: Triticum aestivum

US-10-953-349-26062

Query Match 1.8%; Score 37.8; DB 6; Length 883;
Best Local Similarity 48.4%; Pred. No. 0.29;
Matches 105; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1044 AGAAGGTGGAGAGACTTGAAGTTGCACTCAAGAGGCCAAAGAAAGAGTTTCAGATTTC 1103
DB 558 AGAACCTTGATGTGATGAAAGCTGCTATTGATGGCTTGCAAAACACAGAGGATGATGACG 617
QY 1104 AAAAGAAAACAAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAGCACAGAGAAAG 1163
DB 618 GCAAAGAGATGGTGACACATCTGTGAAGAAAAGCAAGAGAGAAAGTCCAAAGCTGAGG 677
QY 1164 AGAATGATCAAGAGAAAGCCCGGAGACTGTTTGAAGCGAAGTGAAGCACTGAACCTCC 1223
DB 678 CTAACGGTGCAGCGATGGAGCTTGAGATGCTGCAACAGAGCTGAAACTGGAAACCGAGA 737
QY 1224 AGGTGACATCTCTGTTTAAAGGAGCTTCAAGAGCTCA 1260
DB 738 AGAAGAAAGAAAGAAAGCAACAGCTTTGAGGAGCCGCA 774

RESULT 17

US-11-217-529-326
; Sequence 326, Application US/11217529
; Publication No. US20060099612A1

; GENERAL INFORMATION:

; APPLICANT: SUNTORY LIMITED

; APPLICANT: NAKAO, YOSHIHIRO

; APPLICANT: NAKAMURA, NORIHISA

; APPLICANT: KODAMA, YUKIKO

; APPLICANT: FUJIMURA, TOMOKO

; APPLICANT: ASHIKARI, TOSHIHIKO

; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS

; FILE REFERENCE: S-38-285

; CURRENT APPLICATION NUMBER: US/11/217,529

; PRIOR FILING DATE: 2005-09-02

; PRIOR APPLICATION NUMBER: US 10/932,182

; NUMBER OF SEQ ID NOS: 197023

; SOFTWARE: Patent in version 3.3

; SEQ ID NO 326

; LENGTH: 1614

; TYPE: DNA

; ORGANISM: Saccharomyces pastorianus

US-11-217-529-326

Query Match 1.8%; Score 37.2; DB 7; Length 1614;
Best Local Similarity 48.0%; Pred. No. 0.62;
Matches 142; Conservative 0; Mismatches 148; Indels 6; Gaps 1;

Db 1775 CAGGAACAGGAGAGAGATGTGGACGAGGAGGAGAGATGTGTGACGAGGAGAGAG 1834
Qy 545 CAGAAGGAAGAACGCCAGTTTGTGAGATACAGAGCAAGAGCAAGCAAGAGCGTCTAATG 604
Db 1835 ATGCAAGACAGGAGGAGAGATCGGAGGAGGAGGAGAGATGTGTGGAGCAGGAAGT 1894
Qy 605 GCCTTGAGTCATGAGATGAGAAATGAAAGGAGAGCTTGGAAACTAAAG 656
Db 1895 AGGCTGCGCAGCAGGAGGAGAGATGATCAGGAACACCCAGGAGCACCTGGAG 1946

RESULT 21

US-11-217-529-173387
; Sequence 173387, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 173387
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-173387

Query Match 1.8%; Score 36.6; DB 7; Length 849;
Best Local Similarity 49.8%; Pred. No. 0.62;
Matches 120; Conservative 0; Mismatches 119; Indels 2; Gaps 1;
Qy 1503 AAGAACATAATATGCTGAAACAAATTTAGGAACCTAACAGAAAGAGCTCAGAAAAG 1562
Db 390 AGGAACGAGAAATCGTCGGAAGACACAAAGAAAGAACTATAGAGAAAATAAGAGAAA 449
Qy 1563 TGCACAGGGCAGTGTGAGGAACTGAGTGAATA--ACTGGAAGTGCAGAGAGGCTCT 1620
Db 450 AGAAGAGGACGTCGTAAAGCAAGAAAGCAGAAAGTAAAGGAATTTGAAACAGCAGTCGGA 509
Qy 1621 GGCTTCCAAACAGCTGCAAAATGGATGAAATGAAAGCAAAACCATTTGCCAAGCAGGAAGGA 1680
Db 510 GAAGGATAAGAAAGCAAGTTGAAATCTGTTAATAAAGCATACTAACGATCAGAGGA 569
Qy 1681 CCTGGAACCATGACCATCTCAGGCTCAGAGTGAAGTTTACTGTGTTGATTTTCATGC 1740
Db 570 GATCGAAAAGATGCCAAATCAAAGGCTGAATTTGGAGCTTTTAATGATGATGATGA 629
Qy 1741 T 1741
Db 630 T 630

RESULT 22

US-10-953-349-3230
; Sequence 3230, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3230
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-3230

Query Match 1.8%; Score 36.6; DB 6; Length 908;
Best Local Similarity 49.2%; Pred. No. 0.65;
Matches 96; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 467 CAGCTGAAGAGCCATGAAGCTAAATTAATCAAGCCATGAAGGGAGATTGTAGGAGCTT 526
Db 540 CCCAGGAAAGGTGCTAGTGTGGAAGTGTCTTGAAGAGAGAGCCAGCTAGAAGATCTG 599
Qy 527 TCGGCTTGGACAGAGAAACAGAGGAAGAGCCAGCTTTTGGATACAGAGCAAGAA 586
Db 600 AAGAAGCAGAGAGAGATTGCTGCAGCTGCGAAAGAGGATGATAAGAAAAAGAGAGAA 659
Qy 587 GCAAAAGAGCGTCTAATGGCTTTGAGTCATGAGAAATGAAGGAGAGAGCTTGGGA 646
Db 660 GCAAAAGCTCTGCAGCTGTAGATTCAAGCTAAGCTAGACGCTAAGAAAGGTCCAGGC 719
Qy 647 AAATAAAGGGAAA 661
Db 720 AAAGGAAAGGCCAA 734

RESULT 23

US-10-953-349-5081
; Sequence 5081, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5081
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5081

Query Match 1.8%; Score 36.4; DB 6; Length 670;
Best Local Similarity 46.5%; Pred. No. 0.61;
Matches 118; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
Qy 1125 CTGAGATTGAAACCCAGACAGAGGGGAGCAGAGAAAGAGAAATGATGAAGAGAAAGGCC 1184
Db 129 CTGAGAAAGCTGAGAAAGAGGAGAGAGAGGATAAAGATAGAGGATTAAGAACGAGG 188
Qy 1185 CGGAGACTTTGGAGCGAAGTGGAGCACTGAACCTCCAGGTCACATCTCTGTTTAAGG 1244
Db 189 ACGATAAAATGTTGGAGGAGAGAAAGGGGAAAGATCAAGAGAGAGCAAGAGAAAG 248
Qy 1245 AGCTTCAAGAGGCTCATACAAAACCTCAGCGNAGCTGAGCTAATGAAGAGAGACTTCAAG 1304
Db 249 ATAAGAAAGCGAAGAAAGAGAAAGAACCCCTGAAGATAGAAAGATCCAGAGAAAGTTGAAGA 308
Qy 1305 AAAAGTGTCCAGGCCCTTGAAGGAAAAAATTTCTGCAATTCATCAGAGTTTGAATGAAAAAGC 1364
Db 309 TGAAGCTTCAGAAAGATTGAAGAAAAAGATTCAAGCTATGCTTTTAAAGAAAGATGAGATTG 368
Qy 1365 AAGAGCTTTGTTAT 1378
Db 369 TGAAGCTTATTCAT 382

RESULT 24

US-11-217-529-3877
; Sequence 3877, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3877
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-3877

Query Match 1.7%; Score 36.2; DB 7; Length 810;
Best Local Similarity 46.9%; Pred. No. 0.78;
Matches 113; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
QY 1385 AAAAGTTAGCTCAAGTGAAGCATGCTATCAGAAATCAAATGGAACAGGCTAAA 1444
Db |||||
QY 1445 ACAGAGGATCAAAAGTCCAAATTAAGTGTCTACAGATGACACACAAAGCTTCTCAA 1504
Db |||||
QY 1505 GAACATAAATGCAATGAAACAAATGAGGAAGTAAACAGAAAGAGTCAAGAAAGTG 1564
Db |||||
QY 1565 GACAGGCGAGTGTGAAGCACTGAGTGAAGAACTGGAACTGGCAGAGAAAGGCTCTGGCT 1624
Db |||||
QY 1625 T 1625
Db 388 T 388

RESULT 25

US-11-217-529-1760
; Sequence 1760, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1760
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus

US-11-217-529-1760

Query Match 1.7%; Score 36.2; DB 7; Length 1446;
Best Local Similarity 50.3%; Pred. No. 1.1;
Matches 89; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 1501 TCAAGAACATAATATGCAATTGAAACAAATTTGAGGAACATAACAGAAAGAGTCAAGAAA 1560
Db |||||
QY 1561 AGTGACAGGCGAGTCTGAAGGACTGAGTGAAGAACTGGAACTGGCAGAGAGGCTCT 1620
Db |||||
QY 1621 GCGTTCACAAACAGCTGCAATGGAATGAAGCAAAACCATTTGCCAAGCAGGAAGA 1677
Db |||||
QY 1677 GCGTTCACAAACAGCTGCAATGGAATGAAGCAAAACCATTTGCCAAGCAGGAAGA 1739
Db |||||
RESULT 26
US-10-953-349-5386
; Sequence 5386, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5386
; LENGTH: 2730
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5386

Query Match 1.7%; Score 36.2; DB 6; Length 2730;
Best Local Similarity 43.3%; Pred. No. 1.6;
Matches 170; Conservative 0; Mismatches 223; Indels 0; Gaps 0;
QY 1497 TTCTTCAAGAACATAATATGCAATTGAAACAAATTTGAGGAACATAACAGAAAGAGTCAAG 1556
Db |||||
QY 1557 AAAAAGTGGACAGGCGAGTGTGAAGGAACTGAGTGAAGAACTGGAACCTGGCAGAGAAAG 1616
Db |||||
QY 1617 CTCTGGCTTCCAAACAGCTGCAATGGAATGAAGCAAAACCATTTGCCAAGCAGGAAG 1676
Db |||||
QY 1676 GGGATGAAATTAACAGAGTGGAGATGTTGAAGGAAGAGAAAGAGNAGATAGGAAGAAA 1736
Db |||||
QY 1737 AGGACCTGGAAACCACTGACCATCTCTCAGGGCTCAGATGGAAGTTTACTGTTCTGTATTTTC 1796
Db |||||
QY 1797 TGGCAGTTCTGCTGAAAGAGATGATGCTTTTCGAAGACGAGGCGAGGCGAGTCTCTTCATGG 1856
Db |||||
QY 1856 AGATGAGAGAGAGGAGACCAATGATGATAAGAGAGATGAAAAGAGAGAGCAAGGGTT 939
Db |||||
QY 939 AGATGAGAGAGTCTGATGGGCGAGAAACAAGTG 1889
Db |||||
QY 1889 CTAAAAAGCGTGGGAAAGGACGAGTCTCTGGAG 972
Db |||||

RESULT 27

US-10-505-928-344


```
; Sequence 344, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 344
; LENGTH: 4153
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-344

Query Match          1.7%; Score 36.2; DB 6; Length 4153;
Best Local Similarity 50.3%; Pred. No. 2.1;
Matches 89; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1521 TGAACCAATGAGGAACTAACAGAAAGAGTCAGAAAAGTGGACAGGCGAGTGTCTGA 1580
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1342 TTAATGAAATGAAGTTAAGCAGTGGAAAGAAACAAGCGGAAAAAATGAGAAAAGAGCACA 1401

QY 1581 AGGAAGTGAAGTGAAGAACTGGAAGTGGCAGAGAAAGGCTCTGGCTTCCAAACAGCTGCAAA 1640
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1402 ATGAGTACACAGAACTGAGAAAGAAAGAAACGAGATGTGATGTAGACTCAAGAGA 1461

QY 1641 TGAATGAATGAAGCAAAACCATTCGCCAAGCAGGAAGGACCTGGAAACCATGACCA 1697
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1462 AGGAAGAGATGATGCAGACCTTAATAATAAATGAAGAAACTTGAAGAGGAGACTA 1518

RESULT 28
US-10-953-349-20648
; Sequence 20648, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20648
; LENGTH: 1457
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-20648

Query Match          1.7%; Score 36; DB 6; Length 1457;
Best Local Similarity 48.5%; Pred. No. 1.3;
Matches 99; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 995 GAGGAGTTAACTGTGAGCAGCTCTGCTGTGCTTAAGGAAGGGAATCAGAAAGTGGAG 1054
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 787 GTGGAGACAGCTTGGGAAATGATCAAGTCATAGTAAAGGTGTGATTGACCCAGCAAG 846

QY 1055 AGACTTGAAGTTCACCTCAGGAGGCCAAAGAAAGATTCAGATTTTGAAGAAGAAACA 1114
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 847 CTAGTTGATATGTGTACAGAGGACCAAAAACCAAGCTTCTATGTAGTGAAGAGGAGAA 906

QY 1115 AGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGGAGCAGAGAAAGAGATGATGAA 1174
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 907 AAGGAAAGAGCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 966

QY 1175 GAGAAAGGCCCGAGACTGTTGAA 1198
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 967 GAGAAAAAGGAGAGATGTTGAA 990
```

```
RESULT 29
US-10-505-928-758
; Sequence 756, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 758
; LENGTH: 1999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-758

Query Match          1.7%; Score 36; DB 6; Length 1999;
Best Local Similarity 49.0%; Pred. No. 1.6;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1352 TTGAATGAAAGCAAGAGCTTGTATTACTACTACAAAAGTTAGAGCTACAGTGGAAAGC 1411
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1804 TTGTTTTTAAGTAAGACTTTTATTATAACAAAATGTTTTTGGAGCAAAAAA 1863

QY 1412 ATGCTATCAGAAATCAAAATGGAACAGGCTAAACAGAGGATCAAAAGTCCAATTAAC 1471
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1864 AAAAAA 1923

QY 1472 GTGCTACAGATGACACACAAAGCTTCTTCAAGACATAATAATGCAATTCAAAAACAATT 1531
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1924 AAAAAA 1983

QY 1532 GAGGAACTAACAGAA 1547
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1984 AAAAAA 1999
```

```
RESULT 30
US-11-217-529-3292
; Sequence 3292, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3292
; LENGTH: 3279
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-3292

Query Match          1.7%; Score 36; DB 7; Length 3279;
Best Local Similarity 45.4%; Pred. No. 2.1;
Matches 207; Conservative 0; Mismatches 245; Indels 4; Gaps 2;

QY 1353 TGAATGAAAGCAAGAGCTTGTATTACTACTACAAAAGTTAGAGCTACAGTGGAAAGCA 1412
```

Db 995 TGAAGCAAAAGAGAAAGTAAATGATATATTTTGAAGTTGAA--TACCATACGAGATGA 1052
QY 1413 TGCTATCAGAAATCAAAATGGACAGGCTAAACACAGAGATGAAAGTCCAAATTAACGTG 1472
Db 1053 TGTATTATTAAGAAAGAAACTCAAAACGAATATTATAGAGGTAGAACCAAGAAATTCGAAGC 1112
QY 1473 TGCTACAGATGACACACAACAGCTTCTTCAAGAACATATAATGCTATTGAAACAAATG 1532
Db 1113 ACCATCATTTGATACAAAGAGGATTTGGCGAGAAACCCAGAGAAATTCAGAACAACTCA 1172
QY 1533 AGGAATAACAGAAAGAGTCAGAAAAGTGGACAGGCGAGTGTGTAAGGAAGTCAAGTG 1592
Db 1173 ATTGCTGAAAAAGCGTATTTTCAAGATATAGATAACAAAAGAAAGGAATCATCAACAA 1232
QY 1593 AAAAAGTGAATGCGCAGAGAGGCTCTGGCTTCCAAACAGCTGCAAAATGGATGAATGA 1652
Db 1233 AGAAGGAGACCTCCGAGGACTTATTCTGAAATAGACGCAAAAGCAAAATGCCGTTAATCA 1292
QY 1653 AGCAA--ACCATTTGCCAAGCAGGAGGACCTGGAAACCATGACCATCTCCAGGGCTCA 1710
Db 1293 TGAATGAGAGGTATACAGAGCAAGCAGACAGATATAAATTAAGTCCCTCAATCAACTGA 1352
QY 1711 GATGGAAGTTTACTGTTCTGATTTTCATGCTGAAAGAGCAGCGAGAGAGAAAATTCATGA 1770
Db 1353 TAAATTTGTTATTTTAAATCAAGATCAAGATCTGAAAGAGGTTTCTGATGCGGTGCTTTT 1412
QY 1771 GGAAGAGGCAACTGGCATTCAGCTGCGAGTCTTCT 1806
Db 1413 CATTAGGACCATCCCGAAATGAAGGAGAAAATTTCT 1448

RESULT 31
US-10-488-619-1577
; Sequence 1577, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; FILE OF INVENTION: Physiological Conditions, And Genotyping Arrays
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1577
; LENGTH: 641
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-1577

Query Match 1.7%; Score 35.8; DB 6; Length 641;
Best Local Similarity 44.7%; Pred. No. 0.88;
Matches 139; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 1251 AAGAGGCTCATCAAAACTCAGCAGAGCTGAGCTAATGAAGAGAGACTTCAAGAAAGT 1310
Db 326 AAAAAACACAAAACACACAGGGAATAAAAAAACAATAAAAAAAGAAAAAAGAAAAA 385
QY 1311 GTCAGGCCCTTGAAGGAAAAATTTCTGCAATTCATCCATCAGAGTTGAATGAAAGCAAGAGC 1370
Db 386 AAAAAACACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 445
QY 1371 TTGTTTATCTACAAAAGTTAGAGTCAAGTGAAGAGCATGCTATCAGAAATCAAAA 1430
Db 446 CCGAAGAGAAAAAAGCAAGGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 505
QY 1431 TGAACAGGCTAAAAACAGAGGATGAAAGTCCAAATTAATGCTGTACAGATGACACACA 1490
Db 506 AAAAAACACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 565
QY 1491 ACAAGCTTCTTCAAGAACATAATAATGATGATGAAACAAATTTAGGAGAACTTCAAGAAAG 1550

Db 566 AAAAAAAAAAAAAAAAAAGAAAAAACAACAAAAAACAAGAGAAAAAACCACAAAAA 625
QY 1551 AGTCAGAAAAA 1561
Db 626 AAAAAAAAAAAA 636

RESULT 32
US-10-953-349-4498
; Sequence 4498, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4498
; LENGTH: 1319
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-4498

Query Match 1.7%; Score 35.6; DB 6; Length 1319;
Best Local Similarity 47.3%; Pred. No. 1.6;
Matches 107; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 472 GAAAGAAGCCATGAAGCTTAAATAATCAAGCCATGAAGAGGAGATTGAGGAGCTTTCCGC 531
Db 865 GAAAGGAATATTGATGCACGAGTATCGCCGGAGCAAGCGCGTTGGAGGAGCTTAAAGC 924
QY 532 CTGGACACAGAAAACAGAGGAAGAACGCCAGTTTGTGATACAGAGCAAGAGCAAA 591
Db 925 GGAGCTGGAGCGCGGTTCTCGCGTGGCGGTGTTGGTGGGAGAGAGAGGAGGAGGA 984
QY 592 AGAGCGCTCTAATGGCTTTGAGTCATGAGAATGAGAAATGAAAGGAGAGCTTTGGAAACT 651
Db 985 GGAGACACAGAGGAGCTGAGAGAGAGATTGAGATCTGAAAGGTTATTTTGGTAATTT 1044
QY 652 AAAAGGGAATCAGAAAGTCAATCTGAGACCCCACTGATGACTCC 697
Db 1045 AAGGAATGGACAGAGAGTATAGTGGCGCAAAATCGATGATTTCTTC 1090

RESULT 33
US-11-136-524-59
; Sequence 59, Application US/11136524
; Publication No. US20060094034A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: HAREL, Josee
; APPLICANT: BEKAL, Sadjia
; TITLE OF INVENTION: ARRAY AND USES THEREOF
; FILE REFERENCE: 86369-3
; CURRENT APPLICATION NUMBER: US/11/136,524
; CURRENT FILING DATE: 2005-05-25
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 2126
; TYPE: DNA
; ORGANISM: Escherichia coli
US-11-136-524-59

Query Match 1.7%; Score 35.6; DB 7; Length 2126;
Best Local Similarity 52.7%; Pred. No. 2.1;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1758 AGAAAAATTCATGAGGAAAGGAGCACTGGCAATTCAGCTGGCAGTTCTGCTGAAGAGA 1817

Db 1550 ACAAACACGCTGACGGAAGAAAGTGTGTTGGCTGATTAAATGTTCTTCGGGTA 1609
; APPLICATION NUMBER: 10/219,694
; FILING DATE: 14-AUG-2002
Qy 1818 ATGATGCTTTGGAAGACGAGGACGAGTCTCTTGATGGAGATGCGAGTGTGATGGG 1877
; APPLICATION NUMBER: 10/151,081
; FILING DATE: 16-MAY-2002
Db 1610 ATGCTGAATGAGATGCGCCAGAAAGCGGGGTGGAACCTTGAGATCCCTTTACTC 1669
; APPLICATION NUMBER: 10/151,078
; FILING DATE: 16-MAY-2002
Qy 1878 CGAACAAGTGAATCTTGACAGCAG 1903
; APPLICATION NUMBER: 10/125,767
; FILING DATE: 17-APR-2002
Db 1670 CGGAAGAAGTGAAGCCTCTCAGGAG 1695
; APPLICATION NUMBER: 10/287,313
; FILING DATE: 01-NOV-2002

RESULT 34

US-11-293-697-1784
; Sequence 1784, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 1784
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1784

Query Match 1.7%; Score 35.4; DB 7; Length 3283;
Best Local Similarity 59.4%; Pred. No. 3.1;
Matches 60; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
Qy 387 ACCTGGCCACCAAACTGACAGCTTTACCCGGAGGAGCTCTGCGAGCATGAAG 446
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-11-284-877-17

RESULT 35

US-11-284-877-17/c
; Sequence 17, Application US/11284877
; Publication No. US20060095984A1
; GENERAL INFORMATION:
; APPLICANT: Hadlaczky, Gyula
; Szalay, Aladar
; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND METHODS
; FOR PREPARING ARTIFICIAL CHROMOSOMES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 12390 El Camino Real
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/284,877
; FILING DATE: 21-Nov-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 10/808,689
; FILING DATE: 24-MAR-2004

; APPLICATION NUMBER: 10/219,694
; FILING DATE: 14-AUG-2002
; APPLICATION NUMBER: 10/151,081
; FILING DATE: 16-MAY-2002
; APPLICATION NUMBER: 10/151,078
; FILING DATE: 16-MAY-2002
; APPLICATION NUMBER: 10/125,767
; FILING DATE: 17-APR-2002
; APPLICATION NUMBER: 10/287,313
; FILING DATE: 01-NOV-2002
; APPLICATION NUMBER: 09/799,462
; FILING DATE: 05-MAR-2001
; APPLICATION NUMBER: 09/724,872
; FILING DATE: 28-NOV-2000
; APPLICATION NUMBER: 09/724,726
; FILING DATE: 28-NOV-2000
; APPLICATION NUMBER: 09/724,693
; FILING DATE: 28-NOV-2000
; APPLICATION NUMBER: 08/835,682
; FILING DATE: 10-APR-1997
; APPLICATION NUMBER: 08/695,191
; FILING DATE: 07-AUG-1996
; APPLICATION NUMBER: 08/682,080
; FILING DATE: 15-JUL-1996
; APPLICATION NUMBER: 08/629,822
; FILING DATE: 10-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 17084-004018/402Q
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858-678-4777
; TELEFAX: 202-626-7796
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-11-284-877-17
Query Match 1.7%; Score 35.4; DB 7; Length 42999;
Best Local Similarity 61.3%; Pred. No. 15;
Matches 57; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
Qy 1106 AAGAAAACAGTAATCGTTCTGAGATTGAACCCACAGACAGGGGACACAGAAAGAG 1165
; APPLICATION NUMBER: 10/219,694
; FILING DATE: 14-AUG-2002
Db 26842 AAGACAGACAGACACAGTGTGAGAGACAGACAGAGAGAGAGACAGACAGAG 26783
; APPLICATION NUMBER: 10/151,081
; FILING DATE: 16-MAY-2002
Qy 1166 AATGATGAAGAGAAAGCCCGAGACTGTTGGA 1198
; APPLICATION NUMBER: 10/125,767
; FILING DATE: 17-APR-2002
Db 26782 AGAGAGAAAGAGAGACAGACAGATGGA 26750
; APPLICATION NUMBER: 10/287,313
; FILING DATE: 01-NOV-2002

RESULT 36

US-11-217-529-5994
; Sequence 5994, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS

FILE REFERENCE: S-38-285
CURRENT APPLICATION NUMBER: US/11/217,529
CURRENT FILING DATE: 2005-09-02
PRIOR APPLICATION NUMBER: US 10/932,182
PRIOR FILING DATE: 2004-09-02
NUMBER OF SEQ ID NOS: 197023
SOFTWARE: PatentIn version 3.3
SEQ ID NO 5994
LENGTH: 3495
TYPE: DNA
ORGANISM: Saccharomyces pastorianus
US-11-217-529-5994

Query Match 1.7%; Score 35.2; DB 7; Length 3495;
Best Local Similarity 44.2%; Pred. No. 3.7; Indels 2; Gaps 1;
Matches 190; Conservative 0; Mismatches 238;
QY 1086 AAGAGCTTTCAGATTTTGAAGAAACCAAGTAATCGTTCTGAGATTGAAACCCAGACAG 1145
DB 3023 AAGGTGTTTGATGTTTGATAAAGATCCTATTATCATATACGAAGATATTGCGTATGCTA 3082
QY 1146 AGGGGACACAGAGAAGAATGATGAAGAAAGGCCCGGAGACTGTTGGAAGCGAAG 1205
DB 3083 AAGTAGTTTGAGAAAACAAAGAAAGAGGAGAGAGATGAAGAAAGAAAGAAAGAGG 3142
QY 1206 TGGAGCACTGACCTCCAGGTGCACATCTCTGTTTAAAGGAGCTTCAAGAGGCTCATACA 1265
DB 3143 AGAAGAAGAAAGAAAGAGGAAGAGAGAGAAAGAGGAGGAAGAGGAAGA 3202
QY 1266 AACTCAGCGAAGCTGAGCTAATCAAGAAGAGACTTCAAGAAAAGTGTACAGGCCCTTGAAA 1325
DB 3203 AGAAAAGGAGGAAGAGGAAGAAAGAAAGAGGAGGAATAAGAGAGGAGGAAG 3262
QY 1326 GGAAAAATTCGCAATTCATCAGAGTTGAATGAAGAAAGCAAGAGCTGTTTATCTAAC 1385
DB 3263 AGGGGAAAAATAATAAGAGGAAGCAAGAAAGAGGAGGAAGGGAAGAAACAATGAGG 3322
QY 1386 AAAAGTTAGAGCTACAAAGTGGAAAGCATCTATCAGAAATCAAAATG--GAACAGCTAA 1443
DB 3323 AAGAGAGGAGGAAAAATAATAAAGAGGAAGATTAAGAGGAAGAAAGAAAGAAAGAA 3382
QY 1444 AACAGAGGATGAAGTCCAAATTTAACTGTGTCTACAGATGCACACAAACAGCTTCTTCA 1503
DB 3383 GAAAGAGGAAGAGACAGACGGAATCAATGAACACTCTCATTTCCCAAGCTTCCACA 3442
QY 1504 AGACATAAT 1513
DB 3443 ATTACACATT 3452

RESULT 37
US-11-217-529-77456
Sequence 77456, Application US/11217529
Publication No. US20060099612A1
GENERAL INFORMATION:
APPLICANT: SUNTORY LIMITED
APPLICANT: NAKAO, YOSHIHIRO
APPLICANT: NAKAMURA, NORIHISA
APPLICANT: KODAMA, YUKIO
APPLICANT: FUJIMURA, TOMOKO
APPLICANT: ASHIKARI, TOSHIHIKO
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
FILE REFERENCE: S-38-285
CURRENT APPLICATION NUMBER: US/11/217,529
CURRENT FILING DATE: 2005-09-02
PRIOR APPLICATION NUMBER: US 10/932,182
PRIOR FILING DATE: 2004-09-02
NUMBER OF SEQ ID NOS: 197023
SOFTWARE: PatentIn version 3.3
SEQ ID NO 77456
LENGTH: 1278
TYPE: DNA
ORGANISM: Saccharomyces pastorianus

US-11-217-529-77456
Query Match 1.7%; Score 35; DB 7; Length 1278;
Best Local Similarity 45.9%; Pred. No. 2.3; Indels 0; Gaps 0;
Matches 119; Conservative 0; Mismatches 140;
QY 1349 GAGTTGAATGAAAGCAAGAGCTTCTTTATCTAAACAAAAGTTAGAGCTACAAGTGGAA 1408
DB 25 GAGTTGCTTTCCCATCAGAGGAATTTGAAAAGCTTGAAGATCGATGAGATACCTTAAG 84
QY 1409 AGCATGCTATCAGAAATCAAAATGGAAACAGGCTAAACACAGAGATGAAAAGTCCAAATTA 1468
DB 85 AAAAAGAAATCCCAAGAATTGAAGAAAGAAAGAACCAACTATCTGTTACAGCTTCTAATTTG 144
QY 1469 ACTGTGCTACAGATGACACACAAAGCTTCTTCAAGAACATAATATGTCATTGAAAAACA 1528
DB 145 AAAAAGCTTTGAGAGAGAAAGAAAGGCCGATGTTAAGAAAGAAAGTTGCTCGAGATACT 204
QY 1529 ATTGAGGAACTAAACAAAGAAAGAGTCAGAAAAGTGGACAGGCGAGTGTGAAGGAACCTG 1588
DB 205 GAAGATATCAAGCCAGCCCTTTCTAATAAAGAGAAAGAAAGTTGAAAAAGAAATTTG 264
QY 1589 AGTGAATAACTGGAACCTG 1607
DB 265 AAAAATAATGCAAGAACAG 283

RESULT 38
US-10-953-349-19713
Sequence 19713, Application US/10953349
Publication No. US20060107345A1
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nickolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: ENCODED THERBY
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 19713
LENGTH: 1476
TYPE: DNA
ORGANISM: Glycine max
US-10-953-349-19713

Query Match 1.7%; Score 35; DB 6; Length 1476;
Best Local Similarity 57.9%; Pred. No. 2.5; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 45;
QY 1044 AGAAGTGGAGAGACTTGAAGTTGCACTCAAGGAGGCCAAAGAGTTTCAGATTTTG 1103
DB 132 AGAGAAAGAGAGCGCATGATGATGACGTGGAAGCAGAGAGAGAGCTTCAGGTTGGA 191
QY 1104 AAAAGAAAACAAAGTAATCGTTCTGAGATTGAAACCCAGACAGAGGGG 1150
DB 192 AGAAGAAAACGAGACGACGCTCGGAGAGAGAGAGAGAGAGAGGAG 238

RESULT 39
US-10-505-928-326
Sequence 326, Application US/10505928
Publication No. US20060088532A1
GENERAL INFORMATION:
APPLICANT: Ludwig Institute for Cancer Research et al.
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
FILE REFERENCE: 28967/39178
CURRENT APPLICATION NUMBER: US/10/505,928
CURRENT FILING DATE: 2004-08-27
PRIOR APPLICATION NUMBER: US 60/363,019
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 866
SOFTWARE: PatentIn 3.2

```
; SEQ ID NO 326
; LENGTH: 10211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-326

Query Match      1.7%; Score 35; DB 6; Length 10211;
Best Local Similarity 40.9%; Pred. No. 8.1;
Matches 266; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

QY 968 GATGGGCGAAGTATTCTTCGACATCAGGAGTAACTGTGAGCCAGCTCTCTGTGC 1027
DB 7537 GATGAGGCAAAATAATTAATTTGTTTGTCAATCTTCAGTGAATGGCCTCAITCAAGAA 7596

QY 1028 CTAAGGGAAGGAATCAAGAGGTGGAGAGACTTGAAGTTGCATCTCAAGGAGGCCAAAGAA 1087
DB 7597 GTAGAGATGCGACAGCAAACTGGAGRAGAGATGAAGAAATCAGTAGACTGAAAAAT 7656

QY 1088 AGAGTTTCAGATTTTGAAGAAACAAAGTAATCGTTCTGAGATTGAAACCCAGACAGAG 1147
DB 7657 CAAATTCAAGACCAAGAGCAGCTTGTCTCTAAACTGTCCAGGTGCAAGGAGAGCAGCAA 7716

QY 1148 GGGAGCAGAGAGAGAGATGATGACAGAGAAAGGCCGGAGACTGTTGGNAGCGAAGTG 1207
DB 7717 CTTTGGAGGAGCAAACTTAGAACTGAGAAATCTGACAGTGGAAATGGAGCAGAGATC 7776

QY 1208 GAAGCAGCTGAACCTCCAGGTGACATCTCTGTTTAAGGAGCTTCAAGAGGCTCATACAAA 1267
DB 7777 CAAGTGCTACATCAAAATATGCTCTTTGACGACACATTAGAAGTGTGCGAGATTCT 7836

QY 1268 CTCAGCGAAGCTGAGCTAATGAAGAAGAGACTTCAAGAAAAGTGTCAAGGCCCTTTGAAAG 1327
DB 7837 TACAAGAATCTAGAGATGAGCTTGAATTGACAAAATGGACAAAATGCTCTTTGTTGAA 7896

QY 1328 AAAAATTTGCANTTCATCAGAGTTGATGAAAAGCAAGAGCTTTGTTTATACTAACAAA 1387
DB 7897 AAAGTAAACAAATGACTGCAAGAGAACTGAGCTGCAGAGGGAATGCAATGAGATGCCA 7956

QY 1388 AAGTTAGAGCTACAGGTGGAAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACA 1447
DB 7957 CAGAAAACAGCAGAGCTGCAAGAGAACTCAGTGGAGAGAAAATAGGCTAGCTGGAGAG 8016

QY 1448 GAGGATGAAAAGTCCAAATTAATCTGTCTACAGATGACACACAAAGCTTCTTCAAGAA 1507
DB 8017 TTGCAGTACTGTTTGGAGAAATTAAGAGCAGCAAGATCAATTTGAAGGAGCTCACACTA 8076

QY 1508 CATTAATATGCAATTTGAAAACAAATTTGAGGAATCAAGAAAAGAGTCAAGAAAAGTGAC 1567
DB 8077 GAAAATAGTGAATTTCAAGAGAGCGCTAGATTGCTATGCACAAAGACCAGGTGGAAGAA 8136

QY 1568 AGGCGAGTCTGAAGGAGCTGAGTGAAAACCTGGAACCTGGCAGAGAGGCT 1618
DB 8137 GGGAAAGTGAGAGAGGAATAGCTGAATATCAGTACGGCTTCATGAGCT 8187

RESULT 40
US-11-236-238-1
; Sequence 1, Application US/11236238
; Publication No. US200601053641
; GENERAL INFORMATION:
; APPLICANT: Petrukhin, Konstantin
; APPLICANT: Caskey, C. Thomas
; APPLICANT: Metzker, Michael
; APPLICANT: Claes, Wadellus
; TITLE OF INVENTION: BEST'S MACULAR DYSTROPHY GENE
; FILE REFERENCE: 20177P
; CURRENT APPLICATION NUMBER: US/11/236,238
; PRIOR FILING DATE: 2005-09-27
; PRIOR APPLICATION NUMBER: US/09/622,964
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: PCT/US99/03790
; PRIOR FILING DATE: 1999-02-22
; PRIOR APPLICATION NUMBER: 60/112,926

; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/075,941
; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 16125
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(16125)
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 21667, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175,
; LOCATION: 10444, 10445, 10446, 10447, 10448, 10449, 10450, 10451
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10452, 10453, 15338
; OTHER INFORMATION: n = A,T,C or G
US-11-236-238-1

Query Match      1.7%; Score 34.8; DB 7; Length 16125;
Best Local Similarity 58.8%; Pred. No. 12;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 733 CCAGCTCAGACCCAGGTGTGAGGCTACAGGACAGAGAGGAGAGCTGTGGGCATCGT 792
DB 3549 CCAGCTCAGGCGCCAGTGCACCACTCCACTACAACTAAGCTGGGCTCCTGACCAGCTC 3608

QY 793 GTCTGAATGCACTCAAGCTCAAGTCAAGTCCAGCGGCTCTCTCAGA 834
DB 3609 CTGGGCACCTGGAGCTGAGGCTGCGCGCTGGGGGCTGGGCAGA 3650

RESULT 41
US-11-217-529-78785
; Sequence 78785, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78785
; LENGTH: 1992
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-78785

Query Match      1.7%; Score 34.6; DB 7; Length 1992;
Best Local Similarity 58.1%; Pred. No. 3.9;
Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1420 AGAAATCAAAATGGAACAGGCTAAACACAGAGGATGAAAGTCCAAATTAAGTGTGTACA 1479
DB 180 AAAAGTAACTGAATGAAGATGAAGTCCAAATGTAAATTTACGATTAAATTTGGTACT 239

QY 1480 GATGACACACAAAGCTTCTTCAAGAACATAATATGCAATTGAA 1524
DB 240 GACTACTCTACAGAAGCTTGGTACATCCGATAATCTCTGCTTTGTA 284
```

1481 ATGACACACACAAGCTTCTTCAAGAACATAATAATGCATTGAAAACAAATTGAGGGAACCTA 1540

; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 838
; LENGTH: 2784
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-838

Query Match 1.7%; Score 34.4; DB 7; Length 2784;
Best Local Similarity 50.0%; Pred. No. 5.4;
Matches 86; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
Qy 1027 CCTAAGGAAGGGAATCAGAGGTGGAGAGACTTTGAAGTTGCATCTCAAGAGGCCAAGA 1086
Db 632 CCCAGAGAGGGAGAGAGATAACAAGGGAGCAGAGACCAGAGAGAGAAGGGGACAGA 573
Qy 1087 AAGAGTTTCAGATTTGAAAGAAAACAAGTAATCGTTCTCAGATTGAAACCCAGACAGA 1146
Db 572 GACCCAGAGATACAGGGGGCAGAGACCCAGAGACATAGTGGGACAGGGGCCAGAAAGA 513
Qy 1147 GGGGACACAGAGAAGAGAGATGATGAAGAGAAAGGCCCGAGACTGTTGA 1198
Db 512 GGGGACAGAGACCCAGAGAGAAAGGGAGACAGACCCAGAGAGAGAGAA 461

RESULT 46

US-10-953-349-20506
; Sequence 20506, Application US/10953349
; Publication No. US20060107345A1

GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCES-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20506
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Glycine max
US-10-953-349-20506

Query Match 1.6%; Score 34.2; DB 6; Length 1187;
Best Local Similarity 56.8%; Pred. No. 3.7;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
Qy 1578 TGAAGGAAGTGAAGTGAAGAACTGGAAGTGGCAGAGAGAGGCTCTGGCTTCCAAACAGCTGC 1637
Db 32 TGAAGGAGAAAAGATACAGCAAGGATGTGCTGAGGAAGTCATGAATCATTAGATCTGC 91
Qy 1638 AATGGATGAATGAAGCAACACCTTCCCAAGCAGGAAGAGGACCTGGAAA 1688
Db 92 CTATGGATAAATTAAGAAATGTCATGGTGTATCTCTGAAGCAGATGTGGAGA 142

RESULT 47

US-11-217-529-76088
; Sequence 76088, Application US/11217529
; Publication No. US20060099612A1

GENERAL INFORMATION:

; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02

; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76088
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-76088

Query Match 1.6%; Score 34; DB 7; Length 1371;
Best Local Similarity 48.0%; Pred. No. 4.5;
Matches 97; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
Qy 1328 AAAAATTTCTGCAATTCATCAGAGTTGAATGAAAAAGCAAGAGCTTTGTTTACTATAACAAA 1387
Db 892 AAGACTTTTGAACGGCAGACTGGATCANAGTTAAAGGAAGAAGCTGCTTTACTCTATA 951
Qy 1388 AAGTTAGACTCAAGTGGAAAGCATGCTATCAGAAATCAAAATGGAACAGGCTTAAACAA 1447
Db 952 AAAGAAAGCTCAGGAGATGGAAGGAAGTTAGAGAAAAAGAGAAATGAACAAGAAGCTCC 1011
Qy 1448 GAGGATGAAAGTCCAAATTAAGTCTGCTACATGACACACACAGCTTCTTCAGAA 1507
Db 1012 GAATATCGCGAGTTTGAATCCGAAATTAATAAACCTAGAAAGAAACACAAGAAATAAAAGAA 1071
Qy 1508 CATAATAATGCAATTGAAACAA 1529
Db 1072 AATGATACAGTTATGAATACAA 1093

RESULT 48

US-10-511-937-485
; Sequence 485, Application US/10511937
; Publication No. US20060088836A1

GENERAL INFORMATION:

; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 485
; LENGTH: 2165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-485

Query Match 1.6%; Score 34; DB 6; Length 2165;
Best Local Similarity 57.5%; Pred. No. 6;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
Qy 1076 GAGGCCAAAAGAAAGAGTTTCAGATTTTGAAGAAAAAGAAAGTAATCGTCTGAGATTGAA 1135
Db 931 GTGGTGGATGAGAGCCCTCTCTGATTTTGAATACATATTAATGTGTGTGTGATGATCCA 990
Qy 1136 ACCCAGACAGAGGGGAGCAGACAGAAAGAGAAATGATGAAGAGAAAG 1181

Db 991 CCCACACCTGAGGAAGACTCAGAAACACACAGCCTGTGATGAGGAGTAAG 1036

RESULT 49

US-11-217-529-5434
; Sequence 5434, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5434
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-5434

Query Match 1.6%; Score 34; DB 7; Length 2349;
Best Local Similarity 46.9%; Pred. No. 6.3;
Matches 106; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 1345 ATCAGAGTTGTAATGAAAGCAAGAGCTTCTTATACACAAAAGTTAGAGCTACAAGT 1404
|||||
Db 1068 ATCAAGAGCAAAAAGAAATCGATGCTGCTGCTGAAAACAATAGAAAGATACAAGC 1127
|||||
QY 1405 GGAAGCATGCTATCAGAAATCAAAATGGAACAGGCTAAAAACAGAGGATGAAAAGTCCAA 1464
|||||
Db 1128 TTACTGAACGTTCAAGATCAAAATGAAGCAAGGCTCATCTAATAATTGAAAACGCCTC 1187
|||||
QY 1465 ATTAAGTGTGTACAGATGACACACAAAGCTTCTTCAAGAACATAATAATGCAATTGAA 1524
|||||
Db 1188 CTTAATAGAGGAAGTGAAGCTGCTGCTCAAGGTTTAAATTGATCAACAAATGGATTGGAG 1247
|||||
QY 1525 AACAAATTGAGGAAGTAAAGAAAGAGTCAAGAAAAGTGGACAGG 1570
|||||
Db 1248 CACTATTGAAACTAATAAAGTGAACAGAAAAGGGAATATG 1293
|||||

RESULT 50

US-10-953-349-39192
; Sequence 39192, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39192
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-39192

Query Match 1.6%; Score 33.8; DB 6; Length 987;
Best Local Similarity 49.2%; Pred. No. 4.2;
Matches 89; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 459 AGAACACCAGCTGAAAGAACCCATGAAGCTAATAATCAAGCCATGAAGGAGGATTG 518

Db 393 AGAAACGCAAGAAAGATAGGAAACAAAGACAAAGATAAGAAAGGATAAAAGGCCAAAGAAA 452
|||||
QY 519 AGGAGCTTTTCGGCTGGACAGAGAAACAGAAAGGAAGACGCCAGTTTTTTTGAGATACAGA 578
|||||
Db 453 AGGATGTATCGGAACCAAAAGGTTGCAAAAGGATGAAGATGGTAAAGAAAGAGGCCGATG 512
|||||
QY 579 GCAAAAGAAAGCAAAAGAGCGTCTTAATGGCCTTGAAGTCAAGAGGAGGACCAAGAAAAGAAAGAGAGAGGAAAC 638
|||||
Db 513 GCGATGAAGACGAAGAGAGTAAAGAGGAGGACCAAGAAAAGAAAGAGAGAGAGGAAAC 572
|||||
QY 639 A 639
Db 573 A 573

Search completed: May 30, 2006, 00:04:02
Job time : 42 secs